Universidad Autónoma de Madrid
Facultad de Ciencias
Departamento de Biología Molecular

Horizontal gene transfer in *Thermus thermophilus:* mechanisms and barriers

TESIS DOCTORAL

Alba Blesa Esteban

Madrid 2015

Memoria presentada por Alba Blesa Esteban para optar al título de Doctor en Microbiología por la Universidad Autónoma de Madrid, bajo la dirección del Dr. José Berenguer Carlos, en el Departamento de Biología Molecular.

Madrid, Diciembre de 2015

ACKNOWLEDGEMENTS

Ni el libro cerrado da sabiduría, ni el título por sí solo da maestría.
Refranero español

Dice un proverbio chino que cuando bebas agua, recuerda la fuente, y mi abuela que es de bien nacido (¿cómo se puede *nacer mal*?) ser agradecido, así que a ello voy, a la parte más "agradecida" de la Tesis.

Es de rigor que comience agradeciendo al Jefe, Pepe Berenguer, por darme la oportunidad de realizar este trabajo, aunque inesperada. Me siento muy agradecida no solo de sus enseñanzas, sino también de su ilusión por la ciencia. Por su optimismo contagioso, su calidad personal, por tener siempre la bombilla encendida cuando se fundía la mía y por el tiempo dedicado, te doy las gracias.

Quiero agradecer de manera especial a Aurelio Hidalgo, del que no sólo admiro su entrega y rigor científico, sino que sea el ejemplo de las cosas bien hechas. Gracias también por sus lecciones *asertivas* y su apoyo, espalda con espalda.

No sólo se aprende *de*, sino también *con*. Toda la tropa Berenguiana, de pura cepa y de adopción, pasados y presentes, han sido unos compañeros y maestros de excepción, ellos ya saben toda la emoción que esto comporta. Gracias a los ex-combatientes, que me acogieron en su seno, facilitando enormemente mis primeros pasos termófilos. Laura, Bricius (te sigo debiendo una muy grande), Eloy, Leti, Akbar, Yamal, Ángel, Marcos, C. Bayón, Carol, Noé, Mariajo, ¡os echo tanto de menos!. Gracias por la ayuda y los truquillos protocolarios, y, por supuesto, por las risas, las juergas, las cerves, el buen rollo, la procrastinación, el glutamato, e incorporarme a la familia del 109. También quiero agradecer a la nueva generación; Ana, Merce, Nieves (gracias gracias), Mario, Bea, San, Nacho, Marta, muchas gracias por estar ahí para echarme un cable y hacer que no desfalleciera en las últimas pinceladas de la tesis. Y a las constantes, M. Luisa y Esther, por ser unas matriarcas de primera división, dispuestas a ayudar en lo que hiciese falta. Esther, sin ti esto hubiese sido un absoluto caos.

También quiero agradecer a Juan Ayala y Miguel de Pedro, por ser unos vecinos espléndidos y fuentes de sabiduría *old style*. A Ásun y los Amils, por compartir gustosamente los reactivos más extraños. A Dionisio por su disposición colaborativa, a Paco *arregla-todo* y a Ángel por haberme(nos) sacado de más de un atolladero informático. A los servicios de Proteómica, del SMOC, y del SME, por la ayuda técnica y la paciencia.

Gracias a las estancias de la beca FPI, he podido enriquecer este trabajo a la vez que mi inquietud viajera. *Danke schön* Beate Averhoff, for your (and Volker's) warm welcome (*barbequeing* was awesome) and the scientific contribution you've put on my work. I really appreciate all the department's people help but I should highlight Ralf, to whom I feel really thankful for his endless help in and out of the lab, for being the perfect *Thermus* partner beyond the Pyrenees. I brought back home from Frankfurt a bunch of great memories and a large collection of mutants!. También, merci beaucoup Didier Mazel de me donner l'occasion de respirer le *cutting-edge* science à l'Institut Pasteur, *le berceau de la microbiologie*. Grâce aussi à tout le personnel du Mazel's laboratoire, qui sont d'un qualité scientifique et humain exceptionnel. En especial, quería agradecer a Joselito, Rocío y Alfonso (sois unos fenómenos) y al resto de la colonia hispana que hicieron de mi estancia gala una experiencia enriquecedora en todos los sentidos.

En este camino tesístico, el apoyo de la familia, carnal y voluntaria, ha sido fundamental y por tanto, meritoria de mi más cariñosa gratitud. Sin orden pero con concierto, agradezco enormemente los sabios consejos de Geo y Tati, que trascienden más allá de la calle Filadors. A San y Dina, por la amistad surgida de la coyuntura fortuita y a Pari, Charlie, Rodri, Bruno, Sergio, Pedro, Gon, Javi, que me acogieron hace más de una década y han seguido apoyándome con gran cariño. No quiero olvidar a Alberto, Fer, Dani y Juan, siempre atentos a mis avances (y retrocesos) científicos. A las chicas IWQA, las *masters of the Daphnia universe*, por compartir una experiencia nórdica fantástica y seguir ahí de vuelta. A mis amigos del cole, Anto, Raquel, Dani, Ale,... que tras veintimuchos años seguimos siendo una piña fantástica. A mis amigos de toda la vida, Jesús, Nacho, Arantxa, Abel, Bashir, Sergio, Iris, Cecilia, Irene, Juan, Carlos y Amanda, conocedores de mis virtudes y mis defectos, son una fuente de energía, cariño y apoyo; gracias por perdonar mis ausencias, por celebrar con exaltación cada éxito y no dejarme caer en el desánimo.

Igualmente, todos los Novo han contribuido al apoyo y aliento necesario para la culminación de este trabajo. Gracias por compartir las prolongadas comidas, las caminatas domingueras (en 2016 vuelvo, prometido), por la escalada, las cuevas, la nieve, los botillos (y las variantes del mismo) y por enseñarme una filosofía de vida familiar entrañable.

Sin lugar a dudas, el mérito que pueda de este trabajo se lo debo principalmente a mi familia y cualquier defecto, en todo caso, corresponde única y exclusivamente a su autor. Mis padres y hermanos, mecenas de apoyo ilimitado y amor incondicional, son el ejemplo del sacrificio, la honradez, perseverancia y cultura del esfuerzo; una fuente inagotable de comprensión y fuerza. Gracias por haberme enseñado a perseguir las metas sin cuestionarlas y guiarme en el buen hacer. A mis abuelos, tíos y primos, por su alegría infatigable y su interés. Abuela, siento haberme saltado los aperitivos dominicales por ir a cuidar a esos bichos que dices tú. Finalmente, en estos momentos quiero recordar a mi abuelo Ginés, el promotor de mi curiosidad por la ciencia y el que espero que, allá donde esté, se sienta orgulloso del legado que ha dejado.

Y gracias a Sergio, obviamente, que ha vivido y sufrido esta Tesis en primera persona; la mitad de este trabajo te pertenece y has impulsado e inspirado la elaboración de la otra mitad. Gracias por no desistir y luchar, por una nueva esperanza. La gratitud es tan absoluta que las palabras sobran.

Quiero agradecer a lasociedad que a través de la beca FPI que se me otorgó, hizo posible que me dedicara en exclusividad a la realización de este trabajo. Igualmente, agradezco a la SEBBM y la SEM por la financiación para la asistencia a congresos que indudablemente enriquecen el trabajo realizado.

Por fin, amigos, me llena de orgullo y satisfacción, dar respuesta a vuestra pregunta del millón; después de 4 años, !*Habemus Tesis*!

A Teresa y Chito,
por la enseñanza más importante

Table of contents

TABLE OF CONTENTS

LIST OF TABLES

TABLES OF ANNEXES

LIST OF FIGURES

FIGURES OF INTRODUCTION

FIGURES OF CHAPTER 3: MATERIALS AND METHODS

FIGURES OF CHAPTER 4: HORIZONTAL TRANSFER OF VESICLE-PROTECTED eDNA AMONG *Thermus* spp

FIGURES OF CHAPTER 5: CONJUGATION IN *Thermus thermophilus*

FIGURES OF CHAPTER 6: EXPLORATION OF DNA PUSHING DURING *Thermus thermophilus* CONJUGATION.

FIGURES OF CHAPTER 7: PUTATIVE ROLE OF INSERTION SEQUENCES IN CONJUGATION AND DEVELOPMENT OF A TRANSPOSON-BASED MUTAGENESIS SYSTEM IN *Thermus thermophilus*

FIGURES OF CHAPTER 8: DISCRIMINATIVE INTERFERENCE MEDIATED BY THE ARGONAUTE PROTEIN IN *Thermus thermophilus* HORIZONTAL GENE TRANSFER

FIGURES OF CHAPTER 9: DISCUSSION

FIGURES OF ANNEXES

Abbreviations

ABBREVIATIONS

β-gal	β-galactosidase
aa	Amino acids
Abs	Absorbance
ADP	Adenosine diphosphate
Ago	Argonaute protein
Amp^R	Ampicillin resistant
ANOVA	Analysis of variance
ATP	Adenosine triphosphate
BLAST	Basic local alignment search tool
BLASTp	Protein basic local alignment search tool
$Bleo^R$	Bleomycin resistant
bp	Base pair
BSA	Bovine serum albumin
CP	Coupling protein
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxynucleotide triphosphate
DR	Direct repeat
eDNA	Extracellular DNA
EDTA	Ethylenediaminetetraacetic acid
EtOH	Ethanol
EVs	Extracellular vesicles
HEPES	4-(2-hydroxyethyl)1-piperazineethanesulfonic acid
HGT	Horizontal gene transfer
hph5	*hph* thermostable variant of the gene conferring resistance to Hygromycin B
HTH	Helix-turn-helix DNA binding motif
Hyg	Hygromycin
Hyg^R	Hygromycin resistant
IPTG	Isopropyl β-d-1-thiogalactopyranoside
IR	Inverted repeat
IS	Insertion sequence
Kan	Kanamycin

Kan^R	Kanamycin resistant
KDa	Kilodalton
LB	Luria-Bertani culture medium
NCE	Nitrate respiration conjugative element
NI-NTA	Nickel-charged affinity resin
OD_λ	Optical density at λ nm wavelength
OMV	Outer membrane vesicle
ORF	Open reading frame
pAgo	Prokaryotic Argonaute protein
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PFA	Paraformaldehyde
RNA	Ribonucleic acid
RNAi	RNA interference
RNase	Ribonuclease
rpm	Revolutions per minute
SD	Standard deviation
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl polyacrylamide gel electrophoresis
sGFP	Superfolder green fluorescent protein
SOC	Super optimal broth with catabolite repression
sRNA	Small RNA
sYFP	Superfolder yellow fluorescent protein
TXSS	Type X secretion system
TAE	Tris-acetate buffer
TB	*Thermus* broth (rich culture medium)
TEM	Transmission electron microscopy
T-DNA	Transferred DNA strand in conjugation
TNpase	Transposase
TRIS	Tris-(hydroxymethyl)-aminomethane
UV	Ultraviolet light
X-gal	5-bromo-4-chloro-3-indolyl-beta-d-galacto-pyranoside

Presentación

PRESENTACIÓN

La transferencia horizontal de genes (HGT) constituye un elemento central en la evolución de los procariotas, siendo la principal vía de adquisición de atributos relacionados con la resistencia a antibióticos o la adaptación eficaz a ambientes cambiantes, entre otros. Transducción, transformación y conjugación son los mecanismos básicos y fundamentales que permiten el intercambio génico por HGT.

En el género *Thermus* se han descrito altas tasas de transferencia lateral de ADN a través de un sistema de competencia altamente eficiente, así como mediante un mecanismo similar a la conjugación. Este proceso conjugativo, si bien requiere contacto celular y es resistente a DNAsas, resulta ser poco convencional ya que funciona bidireccionalmente, permitiendo el intercambio de ADN entre células isogénicas en ambas direcciones y no depende de ningún mecanismo tipo IV de secreción.. Aunque cualquier *loci* puede ser transferido, el sistema muestra preferencia por genes asociados al megaplásmido pTT27 frente a las de genes cromosómicos (10x), a pesar de estar en el mismo número de copias. Este mecanismo conjugativo permite la transferencia de grandes segmentos de ADN sin ningún orden aparente, al menos en el megaplásmido. Sin embargo, si se detectan zonas del cromosoma más proclives a la transferencia, varias de ellas próximas a secuencias de restricción tipo II que podrían funcionar como OriTs.

Mediante mutagénesis se ha comprobado que los componentes del aparato de competencia natural son necesarios para la conjugación en la célula receptora, pero no en la célula donadora, lo que indica la existencia de dos mecanismos activos independientes en el proceso, uno de bombeo de ADN (*push*) y otro de captura (*pull*) mediado por el sistema de competencia. Las ATPasas hexaméricas funcionales CptA y HerA, parálogas a las helicasas HerA y FtsK en otras bacterias, han sido identificadas como fundamentales en el bombeo del ADN desde la célula donadora.

Así, *T. thermophilus* es capaz de adquirir ADN de cualquier naturaleza, incluso si va encapsulado en vesículas, y por conjugación, con tasas incluso superiores a las registradas en competencia, favoreciendo la adquisición de potenciales ventajas adaptativas que éste pueda otorgarle en un determinado ambiente. Estos mecanismos de HGT son compatibles una batería de estrategias de protección frente a la invasión de parásitos genéticos, como virus, plásmidos y transposones. Entre ellos, la proteína Argonauta funciona como sistema de interferencia DNA-DNA, capaz de proteger a la bacteria frente DNA adquirido por competencia natural. Experimentos *in vivo* demostraron la activación del sistema de vigilancia mediado por Argonauta ante

cualquier sustrato de ADN, si bien, Argonauta no interfiere cuando la entrada de un ADN idéntico se realiza por conjugación, sugirendo que la proteína Argonauta es capaz de discriminar el ADN según la vía de entrada. Así, hipotetizamos que Argonauta actúa frente a un ADN potencialmente peligroso que entra desnudo del medio ambiente, destruyéndolo, y permite la entrada de un ADN "de confianza" adquirido por conjugación. Por ello y por la mayor eficiencia de transferencia en conjugación en comparación con la transformación ante el mismo sustrato, proponemos que la conjugación es la principal vía de intercambio de ventajas adaptativas entre poblaciones de *T. thermophilus.*

.

Summary

SUMMARY

Comparative genome analysis have evidenced that fluent genetic exchange is mainly mediated by Horizontal Gene Transfer (HGT), thus recognized as a leading force of prokaryotic evolution and microbial genetic diversity. Transduction, transformation and conjugation are the three standard mechanisms driving HGT.

The ancient thermophilic bacteria *Thermus thermophilus* laterally transfers DNA in a really efficient way due to a highly sophisticated competence machinery as well as by a conjugation-like process. However, conjugation in *T. thermophilus* emerges rather unconventional as it is efficient between isogenic cells, thus, bidirectional, and no type IV secretion systems seems to be involved. Besides, genes associated to the megaplasmid are transferred with higher frequencies (~10 fold) than those localized in the chromosome. No evident order in the transfer among megaplasmid markers contrasts with the identification of multiple hotspots of transfer in the chromosome. Several of these loci showing higher transfer frequencies are encoded near putative *Tth111* type II restriction sites which might work as OriTs. The competence machinery actively participates in conjugation, being required in the receptor cell but not in the donor. Hence, a two-step model (*push-pull*) is proposed, where the donor energetically pumps DNA to a receptor cell, which actively pulls in the DNA transferred with its competence apparatus. Two active hexameric ATPases, named CptA and HerA, paralogs to the helicases HerA and Ftsk from other bacteria, have been identified as chief components of the pushing step. Therefore, *T. thermophilus* is able to proficiently exchange DNA by transformation (including DNA-protected vesicles which could work as long distance vehicles), and largely by conjugation. This fruitful dynamic gene flow is compatible with a battery of protective strategies to prevent potentially harmful invasion of genetic parasites. Among them, the Argonaute protein elicits a DNA-DNA interference on DNA taken up by transformation. We proved larger insights of Argonaute-mediated interference*,* acting, *in vivo,* against virtually any kind of DNA template. However, when such DNA is transferred by a conjugation-like process, Argonaute is not activated, suggesting a selective immunity towards the way in which DNA is acquired. This fact, together with the higher efficiency shown by conjugation compared to natural competence when the same DNA was transferred, enforcedthe proposal of conjugation as the major motor of shared traits among populations of *Thermus* spp in thermal environments. Finally, we analyse the role of the PrimPol polymerase as a ssDNA guide supplier for the Argonaute protein.

Preface and outline

PREFACE AND OUTLINE

Horizontal gene transfer (HGT) is one of the main forces driving prokaryotic evolution and rapid adaptation to environmental changes. In this Thesis, we have explored how genetic material is laterally exchanged between thermophilic bacteria from the ancient genus *Thermus*, concentrating our study on *Thermus thermophilus*. In particular, we have focused our study on how *T. thermophilus* transfers DNA via a conjugation-like process, aiming to disentangle the mechanism under such proficient transfer. Besides, previous knowledge on a highly efficient transformation apparatus empowered further inspection on HGT processes upholding the genetic transfer as well as the existence of molecular barriers averting it.

Chapter 1 provides an overview of known pathways of HGT in Prokaryotes, emphasizing in the diversity of conjugation systems already depicted. Likewise, we introduce the biological model employed in this research, *T. thermophilus*, its competence apparatus and the recently revealed function of Argonaute protein, displaying a blocking role to DNA uptake in *T. thermophilus*.

Chapter 2 briefly announces the scope of this research and reveals the main objectives of this dissertation.

Chapter 3 includes the materials and methodologies implemented in the experiments foreseen in this dissertation.

In **chapter 4**, we delve into the incidence and transfer of extracellular DNA. We describe how *T. thermophilus* produces extracellular DNA protected in membrane vesicles from degradation by nucleases. Characterization of these vesicles, the DNA enclosed within and its potential role as long-distance and long-term DNA vehicles are discussed.

Chapter 5 encloses a detailed report on the conjugation-like process occurring in *T. thermophilus*. A characterization of the DNA flow by this process, including the preferences and requirements for DNA transfer, are surveyed in order to untie the main players accounting for the genetic transfer. Using a variety of techniques, it is shown

that the competence machinery is involved in conjugal transfer and a two step model (*push-pull*) is proposed.

In **chapter 6**, we provide the molecular basis of one of the main molecules involved in the push-step of the conjugation process. A Ftsk-HerA ATPase showed a central role in DNA donation during conjugation without altering competence. A second ATPase is also involved in DNA donation during conjugation. Characterization of these proteins is accomplished.

Over-representation of insertion sequences across *T. thermophilus* genome in addition to the occurrence of DNA transfer independent from RecA challenged the study of possible site-specific integration through transposases. **Chapter 7** presents a study on transposition signatures and insertion sequences that would provide insights on the genetic rearrangements tailoring *Thermus* genomes, allowing also the design of a transposon-based mutagenesis system.

Chapter 8 concerns on the genetic barriers against the DNA exchanged among *Thermus* cells. We discuss the discriminatory activation of *T. thermophilus* Argonaute protein towards DNA uptake, recently published to mediate a DNA-DNA interference when DNA enters the cell by competence. We show that this system is not activated when such DNA is transferred by conjugation.

Chapter 9 includes an all-embracing discussion about the results depicted in the aforementioned chapters. Besides, recent developments achieved as well as ongoing and future research are remarked.

Major conclusions from this dissertation are listed in Spanish and English.

Chapter 1

Introduction

CHAPTER 1: INTRODUCTION

1.1. HORIZONTAL GENE TRANSFER

Prevalence of a remarkable flexibility among Prokaryotes outcomes from constant genome adjustments associated to punctual mutations, gene duplications and losses, transpositions and other genomic rearrangements. Foremost, prokaryotic genomic diversification, coupled to rapid evolution, are enforced by the fluent acquisition of new genetic information from coexistent organisms by several processes known as horizontal (or lateral) gene transfer (HGT)[8,87,111,149,198]. Moreover, rapid gene exchange *via* HGT confers the ability to rapidly adapt to broad environmental changes[147] which may incur in colonization of new, previously restricted, ecological niches if the HGT genes inherited provide the selective skills for such settlement.

Whole-genome comparative studies verify the impact of HGT on bacterial speciation, which would explain the extensive internal variations among strains within the same (or closely-related) species[28,83]. Undoubtedly, HGT has challenged traditional evolution concepts, but also understanding of these lateral genetic exchanges have empowered comprehension of the spread of resistance genes and pathogenicity-related vectors[146]. Likewise, boost of new formulas to tackle with the understanding of the diversity of gene repertoires have empowered the development of novel genetic tools and molecular protocols which rely on these HGT mechanisms[105,146]. Certainly, its potential biotechnological applications, in addition to the insights provided on evolution and speciation, places HGT in the realm of current molecular biology[121].

1.1.1. Mechanisms of Horizontal Gene Transfer

Three classical pathways drive HGT: transduction, transformation and conjugation (Fig. 1.1)[51,87,119,146]. Allied, vesicles, gene transfer agents (GTAs), nanopods and nanotubes have been described as non-conventional drivers of alternative routes for HGT.

1.1.1.1. Transduction

Bacterial gene exchange by transduction is mediated by bacteriophages accidentally packaging bacterial DNA into their capsids, which is then transferred by infection to a new recipient cell. Genes can be mobilized by means of specialized or generalized transduction, regarding if few specific genes or any bacterial DNA fragment, respectively, can be transferred[31,34].

Transduction impact transcends bacterial diversity, as multiple viral traits such as virulence genes are concurrently transferred, representing, in some cases, more than the 20 % of the host genome as prophages remnants[35].

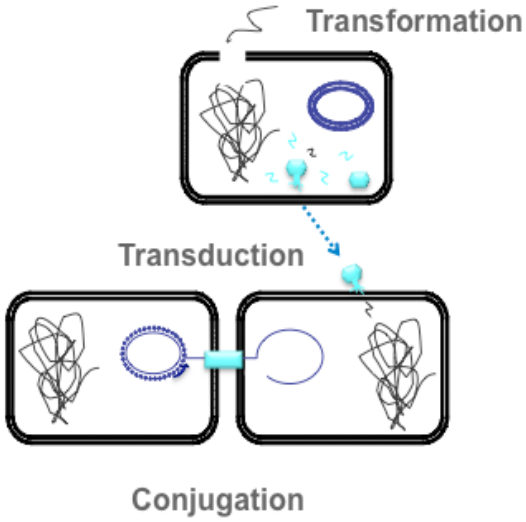


FIGURE 1.1. Main mechanisms of horizontal gene transfer: transformation (uptake of naked DNA), transduction (DNA transfer phage-mediated) and conjugation (cell-to cell DNA transfer).

1.1.1.2. Transformation

Natural transformation involves the uptake, integration and expression of foreign extracellular free DNA, orchestrated by specialized DNA transport machineries[51,74,144,198].

Despite presence of abundant DNA ubiquitously spread in the environmental milieu, just few archaeal and around 82 bacterial species have been identified as competent[107,136]. Nevertheless, potentially transformable bacteria involving novel models of DNA uptake remain yet to be detected[57]. Probably, current underrepresentation of transformation across Prokaryotes may be explained by the constraints of stable transmembrane transport of external DNA, which, in Gram-negative bacteria, requires crossing three barriers (peptidoglycan wall flanked by outer and inner membrane). Besides, transfer involves specialized complex energy-requiring mechanisms, which expression is often dependent on a particular physiological state. Classical transformation mechanisms involve induction of the competence state, binding of the extracellular DNA to the surface of the competent cell, which is then pulled into the cell and transported in a linear fashion to the cytoplasm[74,144]. Then, DNA is translocated into the cytosol ready to be integrated into the host genome by homologous recombination[74,198]. The set of conserved transformation proteins include homologues to type IV pili (T4P) proteins and to proteins involved in type II and IV secretion systems (T2SS and T4SS, respectively) and to ABC transporters (Fig. 1.2).

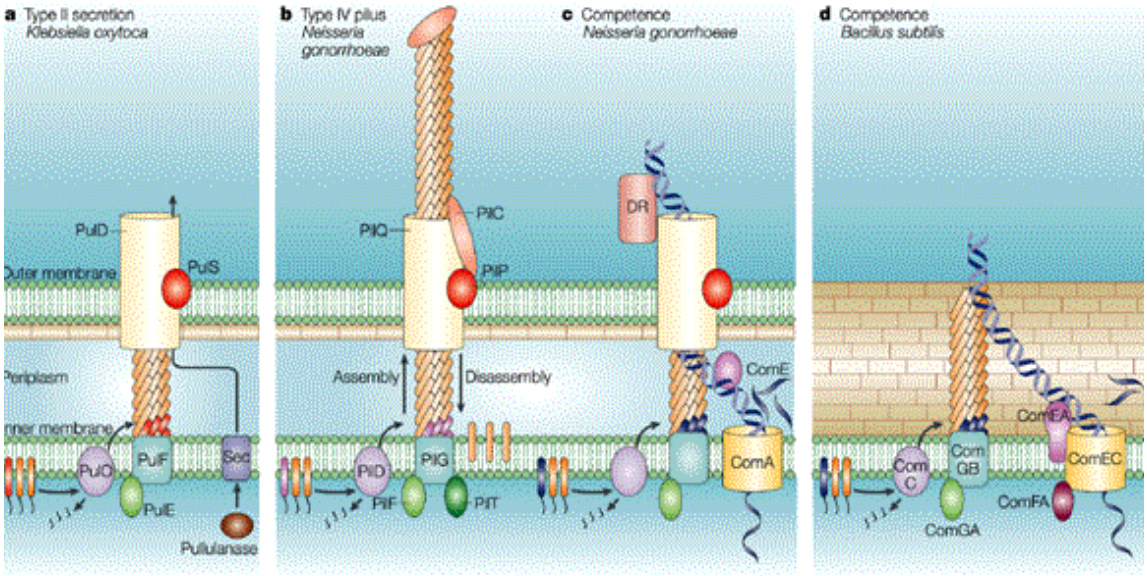


FIGURE 1.2. Key components of the main types of competence machineries. Schematic representation of the most important and differencing components of classical transformation machineries entails: **A.** type II secretion model, based on *Klebsiella oxytoca* pullulanase system; **B.** type IV pilus (T4P) model based on *Neisseria gonorrhoeae* pili; **C.** schematic representation of the pseudopilus model involving DNA uptake and translocator machinery model driving competence in *Neisseria gonorrhoeae*; **D.** Schematic model of competence apparatus in *Bacillus subtilis.* Adapted from Chen and Dubnau[50].

In general, the uptake machinery, composed by homologues of T4P and T2SS proteins, spans the cell envelope, binds to the external DNA and pulls it into the cell, coordinated at the cytoplasmic membrane by ComEC or Rec2 systems[51,50,107,198], where a translocase complex processes it for further recombination with the host genome. Regulation of competence is usually rather sophisticated except for *Helicobacter pylori* (although still disputable) and *T. thermophilus,* for which competence has been described as constitutively expressed. Actually, most natural competent bacteria become temporarily "competent" under specific environmental conditions (*i.e.*, growth conditions, cell density, starvation)[74].

Successful transfer *via* transformation not only relies on the recipient cell needs and capabilities; the nature of the DNA being transferred is also decisive. Apart from the obvious fact that final integration step is easier if DNA taken up has high sequence similarity to the host genome, certain species belonging to Gram-negative families *Pasteurellaceae* and *Neisseriaceae* have shown that efficient uptake occurs only when exposed to specific DNA short sequences, known as DNA uptake sequences (USS and DUS, respectively), which are overrepresented in their genomes[64,75,182].

All in all, transformation represents a chief mechanism of HGT, allowing bacteria to acquire new traits from extracellular DNA. However, it is also important to note that DNA is also a formidable carbon and phosphorous source[53,151], dispensing, thus, the energetic fitness cost of nucleotide synthesis. In this line, several authors

have proposed that acquired DNA by transformation satisfies (in addition or preferentially) a nutritional purpose[78,158]. The so-called "nutritional competence"[153] agrees with the fact that in many microorganisms competence is only induced under starvation[78,184], and that any kind of extracellular DNA may be acquired, including sequences with not enough homology for a recombination event.

1.1.1.3. Conjugation

One of the most widely distributed HGT mechanism is conjugation, a highly specialized process by which DNA is transferred between two cells which are in direct contact. First described by Lederberg and Tatum[193], it embodies a key element in dissemination of virulence genes and antibiotic resistances as well as an important source of genetic plasticity[146,181,198]. Ample literature reflects the broad-host-range of conjugative events, intra- and inter-species, where wide range of bacterial genera and some Archaea dynamically share genetic information, occasionally embodying trans-kingdom transfer[143,154,172].

Classical conjugation entails mainly a plasmid-encoded process, where single-stranded DNA (ssDNA) is unidirectionally transported from an unique *cis*-acting site (*oriT*) in known genetic sequence order, connected to a DNA rolling-circle replication (RCR) system. DNA transport requires in most cases a T4SS[122] building a bridge between the donor and the recipient cell which basically remains passive. In this system, *mob* genes (also known as *dtr*) endow DNA mobilization and *mpf* genes encode the mating pair formation machinery (Fig. 1.3)[32,181]. Localization of these genetic elements discriminates between conjugative and mobilizable plasmids. "Conjugative" plasmids are self-replicating plasmids carrying the complete set of both *mob* and *mpf* genes[181], whereas "mobilizable" plasmids lack *mpf* genes and require the assistance of *mob* genes encoded in other regions of the genome of the donor strain, such as integrative conjugative elements (ICEs), to warrant the transfer[95,213].

Usually, DNA is transferred *via* small-size conjugative plasmids, which are abundant and widespread among bacteria, starring quick transfers according to plasmid mobility principles[181]. However, large sections of DNA can be mobilized too, even the whole chromosome, when these conjugative elements happen to integrate in the chromosome. Regardless of the genetic information transferred, comparative studies have evidenced that there is a strong gene synteny among most bacterial conjugative systems, allowing the development of an integrative two-steps model of conjugation[32].

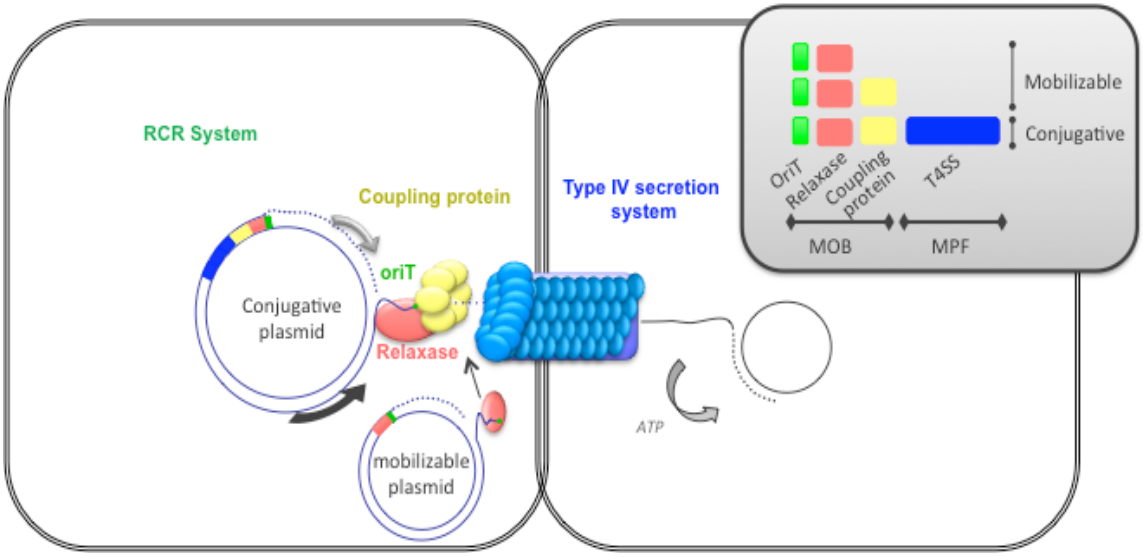


FIGURE 1.3. Conjugative machinery is encoded in *mob* and *mpf* genes and DNA is transferred following a RCR system-like. Schematic model for the classical conjugation process including the main elements managing the transfer (see text). At the upper right corner of the figure, a schematic view of the genetic constitution of transmissible plasmids is provided too, modified from Smillie *et al*[181].

Conjugative molecular machinery

The initial step of conjugation involves the DNA processing, transfer and coupling to the secretion system. It is enforced by proteins coded at the *mob* region which, together with the substrate *oriT* sequence comprise the relaxosome complex[84] (Fig. 1.4). Thus, this nucleo-protein complex includes an origin of transfer (*oriT*), a relaxase, which is a ssDNA transesterase present among all conjugative plasmids, and several accessory proteins involved in DNA cleavage. Additionally, a well conserved ATPase known as coupling protein, outcomes fundamental in the conjugative process as a connector to the secretion apparatus and energizing system.

The conjugative process is initiated by contact between two cells. This contact is mediated by an adhesin located at the distal part of one of donor's external appendages, the conjugative pili[7,9]. More than a dozen proteins are involved in assembling this pilus, which by depolymerisation of the base will bring closer both mates. Meanwhile, the relaxase cleaves a phosphodiester bond of the strand to be transferred in the DNA molecule (T-DNA) at an specific site (*nic* site) in the *oriT*[49]. In some cases, relaxases are bi-functional and fuse to an helicase or primase domains[216]. Auxiliary transfer factors bind to the *oriT* so as to ensure that such cleavage is specific[84]. Simultaneously, the complementary strand is replicated alike a RCR-system. The nucleo-protein comprised by the T-DNA strand covalently bound to the relaxase is recruited to the channel of the T4SS, where it docks to the coupling protein to initiate transfer, energized by its ATPase activity[122]. Once again, auxiliary transfer factors are frequently involved in the binding of the coupling protein to an specific region of the relaxosome. Akin, the second step of conjugation is determined at the *mpf* region

which encodes a T4SS that allows DNA passage to the recipient cell[55] and the T4P involved in the capture of the recipient cell and the approach of both cells.

Despite it is still unclear how the relaxase is transported across the secretion channel, how the T-DNA is actually translocated and the conformational changes taking place is still matter of controversy, several models of DNA translocation have been postulated: the channel model[38], the ping-pong model[13] and the shoot and pump model[122], which differ in the energizing factors and the role of the coupling protein once T-DNA reaches the receptors' membrane. Finally, once a copy of the single strand is completely transferred, the relaxase re-circularizes the transferred strand in the receptor through DNA ligation[72].

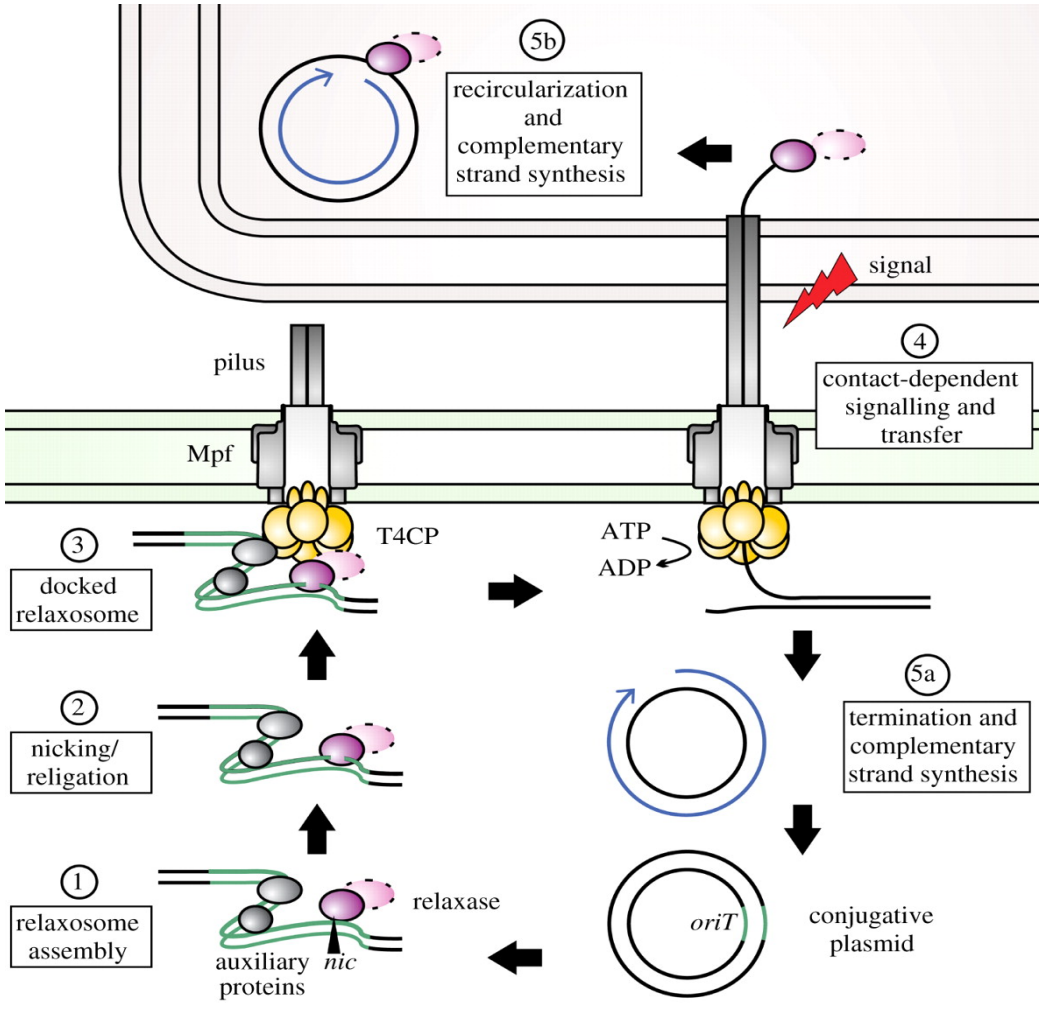


FIGURE 1.4. General mechanism of plasmid transfer by conjugation. *mob* and *mpf* genes harboured in a conjugative plasmid are expressed. Relaxase in purple, (bi-functional relaxases are in pink), the coupling protein in yellow and in grey the translocator apparatus. T-DNA is processed and transferred as explained in the text. Adapted from Zechner *et al*[216].

In certain occasions, the conjugative machinery from conjugative plasmid, is integrated into the host chromosome, either as an Integrative and Conjugative

Elements (ICEs)[95,213] or giving rise to *Hfr* (high-frequency recombination) cells[59,193]. In these cases, ICEs excise, circularize in plasmid-shape before transfer whereas in *Hfr* cells the transfer pulls donor's chromosomal region aside, as single-strand DNA. ICEs re-integration into host's chromosome takes place at a certain site, where they are maintained upon next mobilization whereas in *Hfr* systems transfer is generally interrupted by random breakage, impeding further transfer[46].

Unconventional models of conjugation

While alternative models to traditional conjugation for the cell-to-cell DNA transfer are not abundant, they are nevertheless present and varied in the literature. In Mycobacteria lacking conjugative plasmids, a chromosome-encoded unconventional conjugation system named Distributive Conjugal transfer (DCT)[70,140] has been described. In DCT extensive fragments of non-contiguous chromosomal DNA are transferred from multiple sites in a unidirectional fashion from the donor to the recipient cells, in the absence of known plasmids or ICEs[206]. Indeed, in contrast to *oriT*-mediated transfer, DCT apparently shows multiple *cis*-acting sites as various regions of the chromosome are transferred at similar efficiencies, regardless of its chromosomal location[207]. Hence, multiple initiation sites should exist on the chromosome which would confer transferability in this location-independent manner. Moreover, DCT foresees that some transconjugants can become donors, in contrast to *Hfr* systems, which would require the transfer and recircularization of the whole chromosome[207]. Moreover, DCT provides an outstanding genomic variation upon transconjugant progeny, which genome shows a complex structure, mosaic of the DNA inherited from parental strains[70]. *In silico* approaches have failed to identify any orthologs of the transfer machinery, hindering further untying of the mechanism underlying. Nonetheless, genetic assays have related DCT with genes encoding the ESX-1 secretion apparatus, a pathogenic-related type VII secretion system (T7SS), which would be playing relevant roles mediating DCT, sometimes occasioning an hyper-conjugation phenotype in the donor[79]. Certainly, genus synteny of this conjugative system has accelerated further rethinking of general conjugation model, buoying research for alternative conjugative models, which undoubtedly would profoundly impact on genome dynamics and bacterial evolution knowledge.

Apart from chromosomal DCT, Mycobacteria possess a wide range of particular conjugative plasmids evenly distributed performing highly efficient interspecies transfer[62]. Among these conjugal episomes, some harbour a complex system, involving

type IV and type VII machineries for DNA secretion, mediating an atypical transfer of large fragments of DNA in slow-growing mycobacteria[201].

In addition to these atypical systems, *Streptomyces* exhibit another unconventional conjugative mechanism involving proteins which have had to overcome *Streptomyces* thick and complex cell walls, empowering the interaction with membranes, peptidoglycan and DNA[205]. Actually, most Actinomycetes carry small multi-copy self-transmissible cryptic plasmids which are transferred with high efficiency, in a remarkably different way from that of general models[197]. Dual plasmid-borne and chromosomal transfer is performed by a single plasmid-encoded DNA-translocase, TraB, which mobilizes dsDNA molecules to the recipient upon recognition and binding to repeated 8-bps motifs on the conjugative plasmids[196,197,205] (Fig. 1.5). TraB, localized at the pole of the hyphal tip of the *Streptomyces* cells, and assembles as an hexameric pore-forming ring probably adapted from Ftsk/SpoIIIE systems[160,196].

This protein controls the mobilization of chromosomal genes and mediates intermycelial plasmid transfer. Hence, TraB, assisted by plasmid-encoded Spreading proteins (Spd-proteins), also leads the DNA translocation by a complex apparatus, followed by a rapid spreading of the newly acquired DNA via septal cross walls, attaining a prompt colonization of the recipient mycelium[112,197].

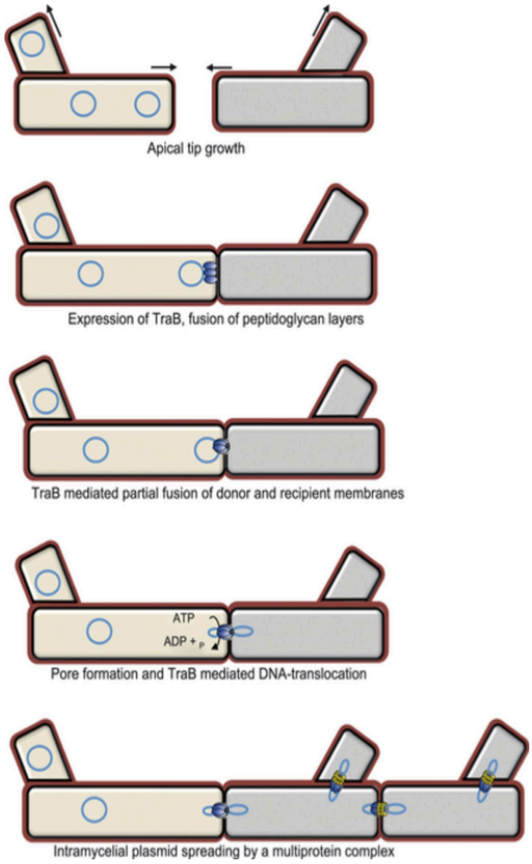


FIGURE 1.5. Conjugative plasmid transfer in *Streptomyces*. The schematic model of plasmid conjugation occurring in *Streptomyces*, orchestrated by TraB, as explained in the text. Adapted from Thoma and Muth[196].

Among Archaea, several reports document conjugation occurring mainly among halophilic and non-methanogenic hyperthermophilic Archaea[125,186]. Unlike standard bacterial conjugation, *Halobacterium* spp and *Haloferax* spp mediate DNA transfer in a bidirectional fashion, exchanging chromosomal genes by hypothetical membrane fusions[142,163]. *Desulfurococcus mobilis* performs intron and chromosomal gene transfer at high rates[1] and *Sulfolobus* spp[93] harbour pNOB8, a self-transmissible conjugative plasmid that exhibits highly efficient unidirectional

transfer and is propagated at high copy number in the recipient cells[156]. Besides, there is cornerstone evidence supporting successful inter-domain conjugal transfer between Archaea and Bacteria[71,143], congruent with the ubiquitous presence of archaeal genes across the three domains of life.

1.1.1.4. Newly discovered routes of HGT

Intense research on prokaryotic gene exchange has compelled the emergence of alternative mechanisms of HGT, working as inter-cell DNA vehicles.

Gene Transfer Agents (GTAs) are bacteriophage-like elements produced by several α-Proteobacteria that package randomly short fragments of dsDNA from the bacterial genome which can be this way transferred at high frequencies[117,134]. Quorum sensing, peptide signalling and a complex combination of environmental factors regulate GTAs expression, encoded by conserved gene clusters[116,135]. Prevalence of GTA-encoding genes, widespread among environmental bacteria, suggest sound relevance upon acquisition of advantageous adaptative traits, thus, potentially displaying an important role on Prokaryotic evolution[116].

Nanopods, a recently discovered class of extracellular appendices, have come to the fore as ancillary HGT tools, working as long-distance DNA vehicles[179]. Found in several members from *Comamonadaceae* family, nanopods consist of strings of outer membrane vesicles (OMVs), sheathed by a tubular assembly of a glycosylated surface layer protein, NpdA. Nanopod-forming OMVs can contain DNA, proteins or signalling substances which can be successfully transferred. Production of these novel bacterial surface structures is driven by OMVs production, which in turn responds to bacterial communication, pathogenesis and nutrition[180]. Nanopods have proved to conduit long-distance OMVs-mediated DNA delivery[179], albeit at low frequencies. Nevertheless, its deployment sets a new paradigm in mechanisms driving long-distance interaction between bacteria.

Nanotubes, similar to nanopods, are extracellular tubular structures formed by membrane proteins protruding from the cell which have evidenced its capability to transport cytoplasmic constituents, including proteins and DNA, between neighbour cells[73]. First reported in *Bacillus subtilis*[73] and then on different microorganisms including hyperthermophilic Archaea[132], these nanotubes would work as intercellular bridges. Hypothetically originated by direct cytoplasmic fusions of adjacent cells,

nanotubes would thereby release cells from expressing a particular secretory machinery for DNA and protein transfer, ultimately facilitating dynamic inter-bacterial communication.

1.2. PROKARYOTIC DEFENCE STRATEGIES

On behalf to so many different ways that genetic material may be fluently exchanged, uncertain exogenous DNA might reach the cells, jeopardizing their welfare. Therefore, microorganisms have developed a battery of multi-layered protective strategies acting against the invasive bombardment of detrimental genetic parasites, mainly viruses, but also plasmids and transposons (Fig. 1.6). Genes encoding defence systems are widely distributed, often clustered in the so-called defence islands, either localized in the chromosome or linked to mobilizable elements, thus, prone to be transferred[130].

Among the classical protective arrangements, we find the restriction-modification systems, which provide an innate immunity response to nuisance DNA incoming. Restriction endonucleases actively cleave non-self dsDNA at specific sequences' motifs, and act in combination with modification enzymes, mainly methyltransferases, which modify host's self DNA, preventing self-cleavage of their own genome[199].

Besides, plethora of studies have documented how prokaryotes defend themselves following programmed cell death or dormancy induced by infection. Numerous toxin-antitoxin (TA) systems as well as abortive infection (Abi) systems have been described[54] as efficient mechanisms mediating blockage of hazardous DNA. Both systems are activated once the infection is sensed, triggering a multi-enzymatic activity often directed to interfere the host's translational processes, thus, detrimental for the cell metabolism, which, in Abi cases, may foster a fatal outcome[130].

Further defensive mechanisms follow ideal tactics to withstand viral predation. Some systems intercept viral infection at the cell membrane, preventing viral attachment (adsorption blocking systems) or injection (uptake blocking systems) by masking host-cell's receptors[115]. If a virus has already been attached to host's surface, superinfection exclusion (Sie) systems can mediate blockage of DNA entry into the host cell. A multi-complex consisting of membrane anchored and membrane associated proteins interact with viral proteins involved in DNA passage, impeding its progress and integration. These mechanisms, albeit lesser known, are of critical relevance for the surrounding population as it renders the incoming phage non-infectious[175]. Recently, Goldfarb *et al*[89] have unveiled a novel defence system against

bacteriophages' DNA replication, named Bacteriophage Exclusion (BREX) systems. The system follows a host DNA methylation-based defence which mechanism is still undetermined. Nonetheless, the authors assert that it is widespread among Prokaryotes and confers protection against lytic and lysogenic phages, although it does not inhibit viral adsorption. Interference of phage production following infection is achieved by limiting phage growth through the expression of a gene-cluster encoding enzymes linked to hosts' DNA methylation.

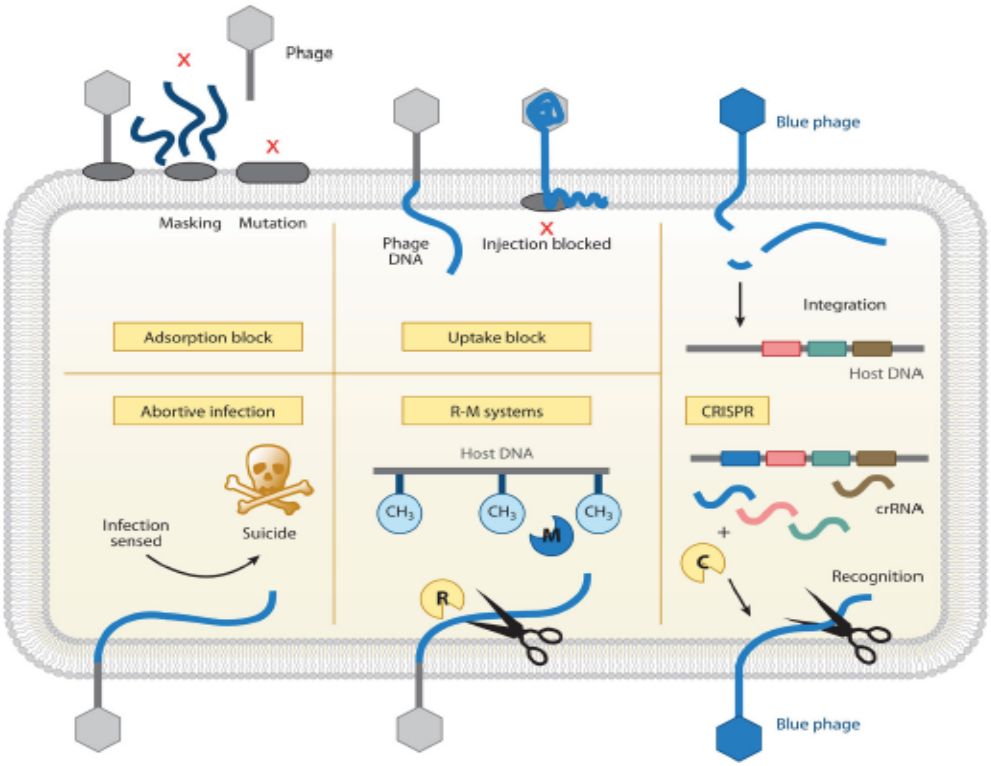


FIGURE 1.6. Mechanisms of prokaryotic defence. Schematic overview of some of the defensive mechanisms that Prokaryotes may exhibit while battling phage infection. Adapted from Makarova *et al*[130].

In addition, Prokaryotes have developed exceptional adaptative immune responses based on the interference of nucleic acids. The nowadays popular CRISPR/Cas system, overseeing RNA-DNA interference and prokaryotic homologues of Argonaute proteins display crucial roles in cell protection.

CRISPR/Cas systems, exclusively found in Prokaryotes, neutralize invaders by a complex mechanism, reminiscent to eukaryotic RNA interference[101,128], based on the incorporation of short invaders' sequences in host-cell genome, which serve as molecular memory preventing further reinfections (Fig. 1.7). Widespread among half of bacterial species and almost every Archaea[92], is the only adaptative and inheritable prokaryotic immune system known to date[211]. It is also worth-mentioning that chronological ordering of spacer enables comparative studies among strains within the same species, disentangling evolutionary changes within a population and throwing

light to the perpetual arms-race between bacteria and phage which harshly shapes the evolution of microorganisms[194].

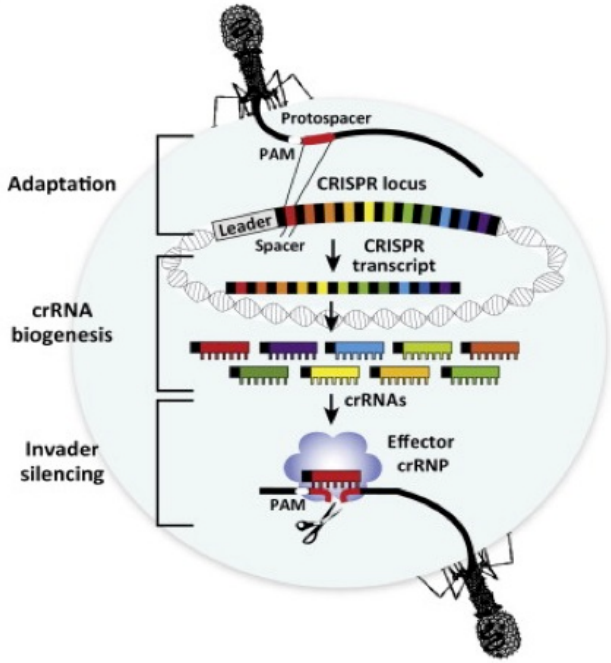


FIGURE 1.7. CRISPR locus and pathways. Coloured invader-derived sequences (spacers) are interspaced by hosts' direct repeats, represented in black. Preceding, the AT-rich leader sequence and the *cas* genes (not shown). Protospacer adjacent motifs (PAMs) are located immediately nearby to spacers selected for CRISPR integration. The three phases entangling a CRISPR-Cas action (adaptation, expression and interference) are illustrated. An initial adaptation phase, involving the acquisition of a new DNA sequence and its integration into the hosts CRISPR locus adjacent to the leader sequence. Then, CRISPR locus are transcribed and these transcripts are processed generating multiple mature CRISPR RNAs (crRNAs), each one specifically targeting a different sequence. Finally, interference is executed by Cas proteins mediating gene silencing by recognition of foreign nucleic acids through base-pairing of crRNA. Adapted from Terns and Terns[194,195].

Despite wide evolutive divergence among CRISPR/Cas systems, undoubtedly promoted by the inner diversity of spacers, a classification of the different mechanisms governing CRISPRs interference has been proposed (Fig. 1.8), embraced in three main groups. Type II, the simplest CRISPR system, encloses Cas9, which plays a pivotal role in the interference, performing endonuclease and integrases activities with ground-breaking accuracy by itself[203]. In fact, the outstanding properties of this enzyme prompted scientists of CRISPR/Cas9 potential role as a gene re-editing tool, recently worldwide awarded.

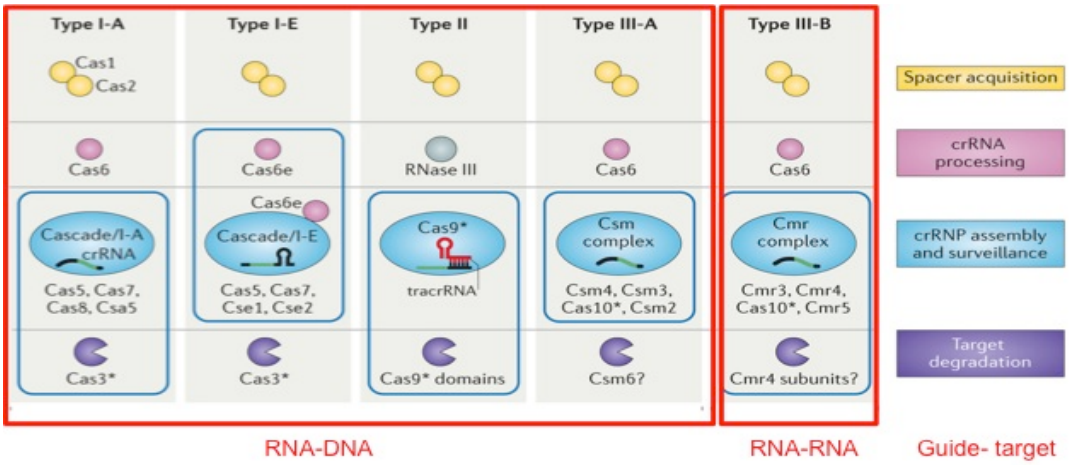


FIGURE 1.8. Diversity of CRISPR-Cas systems. The 3 main types, share a core set of Cas proteins (Cas1-6), and several subtypes are defined within each case, including specific modules such as Csa, Csm, Csd or Csy, among others[203,211]. Typologies are defined on the basis of a specific signature that a certain Cas protein (indicated by an asterisk) displays on each process. Blue boxes enclose CRISPR ribonucleoprotein (crRNP) complexes which, preloaded with the small RNA guides, execute the silencing of the targeted nucleic acids. Modified from van der Oost *et al*[203].

Along the latest years, a novel prokaryotic defence system based on homologues to the eukaryotic Argonaute (Ago) protein has been identified on at least 9 % of bacteria and 32 % of archaea. Indeed, most Ago are highly specialized small-RNA-binding proteins sharing substantial structural similarities throughout all domains of life[30,190]. In Eukaryotes, gene silencing is orchestrated by a complex network of processes involving diverse highly specialized players and factors, guided by small-RNAs (sRNAs) to control gene regulation and expression[106] (Fig. 1.9).

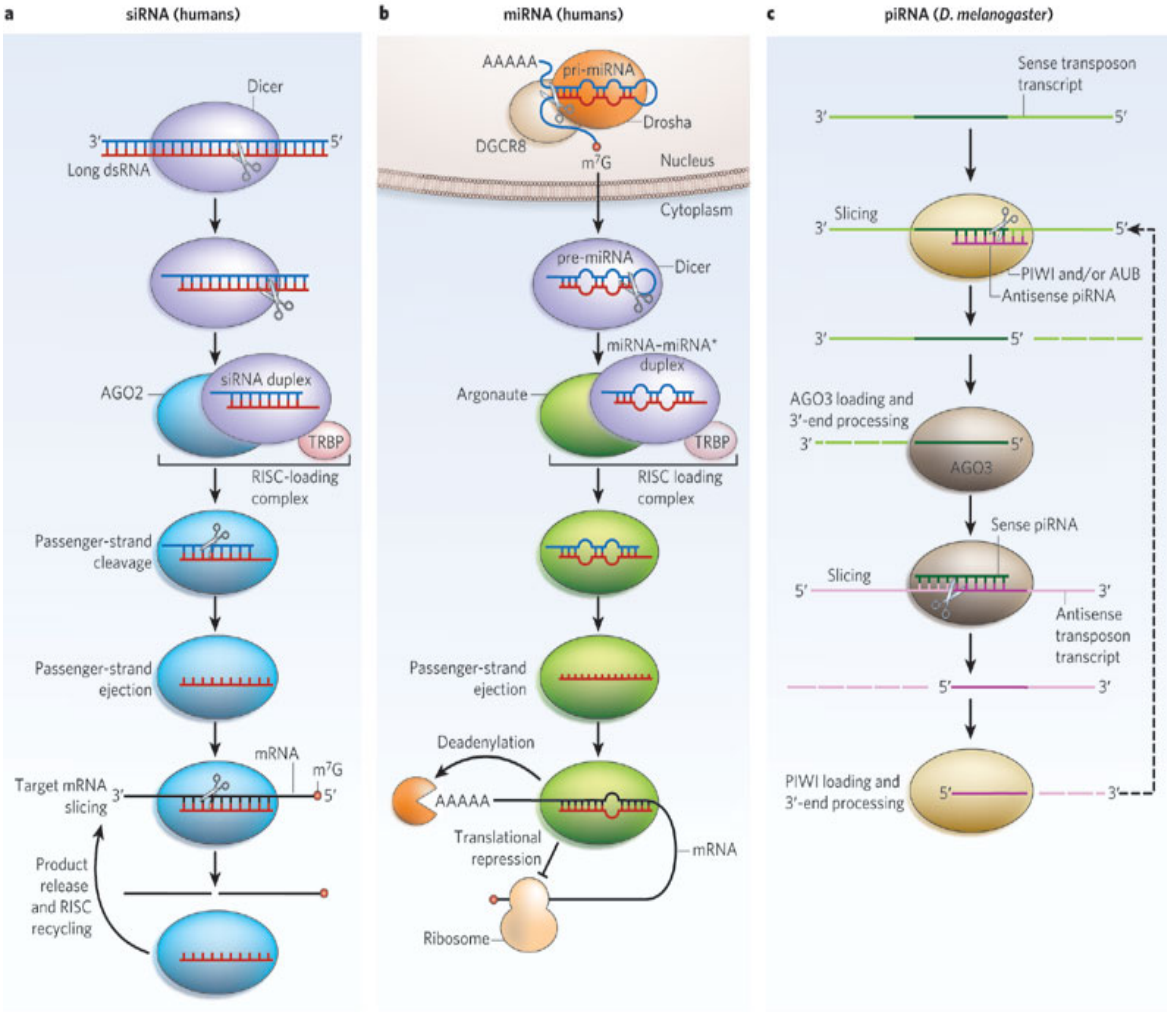


FIGURE 1.9. Classical pathways of small-RNA induced gene silencing. Three kinds of small RNA molecules have been defined in Eukaryotes, headlining the main pathways of small-RNA-guided interference: in the left panel, small interfering RNAs (siRNA), which are generated by Dicer's enzyme cleavage, and tightly bound to Argonaute forming the RISC complex; in the centre, microRNAs (miRNAs) which are twice processed in the nucleus and the cytoplasmic by Drosha and Dicer enzymes, respectively, after which are guided to Argonaute, then interacting with translational machinery and deadenylases; and, in the right panel, the Piwi-interacting RNA (piRNAs), involved in transposon silencing, where two Argonaute proteins (Piwi or Aubergine and Ago 3) are interspersed during piRNA maturation. Extracted from Jinek and Doudna[106].

Ago proteins constitute the functional core component of these gene silencing machineries, as these versatile endo-ribonuclease uses sRNA guide molecules (20-30

nucleotides) to specifically target the complementary RNA target. Hence, Ago embodies the direct junction of the RNA base-pairing interfered[106,191] and enacts cleavage or blockage of the target, resulting in interference of such RNA[47,76,100,106,177]. By these means, Ago proteins may control a wide range of physiological and cellular processes by RNA interference[209].

Ago proteins belong to the PIWI domain superfamily, a highly conserved family encoding diverse regulatory proteins, often featured by a RNase H fold which entails a nuclease activity[100,106,171]. The Ago family, which name was coined according to the resemblance of the Ago-knockout phenotype in *Arabidopsis thaliana* to the tentacles of the octopus *Argonauta argo*[20], comprises large proteins (around 100 Kda), strongly basic (pI: 9-9.5), arranged in four well conserved modules: N-terminal, PAZ, MID and PIWI domains, interconnected by two structured linker regions (Fig. 1.10).

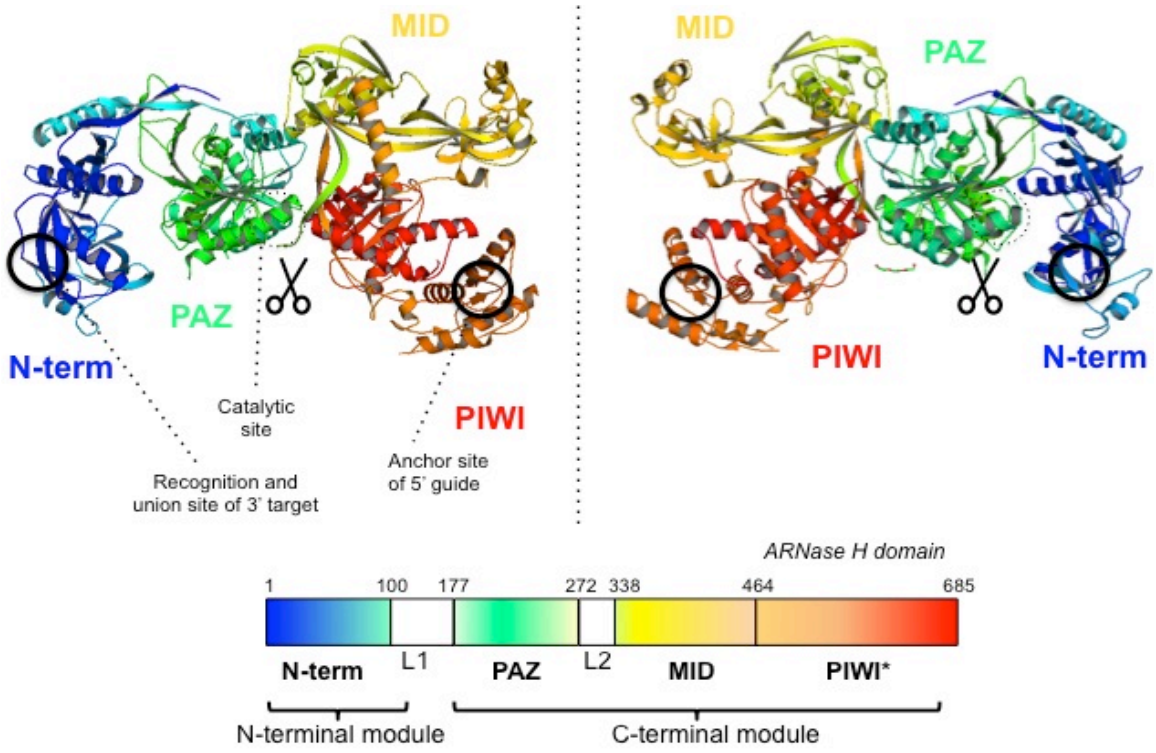


FIGURE 1.10. Ago proteins are comprised into 4 domains with different functions. A highly conserved bilobial architecture entails the 4 domains conforming Ago proteins**.** In blue, the N-terminal domain works as a wedge during dissociation of passenger and cleaved target strands as well as in target cleavage, although mechanistic insights about the latter remain currently unclear[171,190]. In green, the PAZ (Piwi-Argonaute-Zwille) domain has a role in 3'-end recognition and displays a crucial role in guide protection against degradation[208]. In yellow, the MID (middle) domain is in charge of anchoring the small RNAs onto the Ago protein in a helical conformation by specific binding of the 5' phosphates of the sRNAs, upholding, thus, target binding[77]. Finally, in red, the PIWI (P-element-induced wimpy testis) domain which intimately interacts with MID domain, shows a Rnase H-like slicing activity, developed by a catalytic tethrad (*D478,D546, D660 and Glu512)[185,190], represented by scissors and indicated with an asterisk (*). In addition, recognition union and anchor sites are indicated. Figure was prepared using PyMol on the *Thermus thermophilus* HB8 Argonaute template, taken from the PDB database (pdb 3dlh).

The architecture and domain composition is preserved through the four families of eukaryotic Agos: the Ago-like family, the PIWI family, the WAGO family and the

Trypanosoma Ago family[100,191]. According to sequence homology, Ago proteins are clustered into two main types, *Arabidopsis* AGO1 type and *Drosophila* PIWI-like Ago[76]. Besides, subgrouping may rely on their capacity to induce target cleavage or the lack of slicing ability, mediating target silencing by coupled translational repression and deadenylation with mRNA decay[171].

In contrast to the profuse documentation on the biogenesis, evolution and biological role of eukaryotic Ago proteins, scarce information has been available on prokaryotic Agos (pAgos), limited to few structural studies which smoothed the understanding of eukaryotic paralogs[148,178,191,208,212]. *In silico* approaches certify pAgos genetic conservation and synteny despite being subjected to HGT[129], and show that pAgos are often co-localized within operons containing endonucleolytic Dnases and helicases[30]. However, homologues to several proteins essential in (most) eukaryotic RNAi pathways are absent in prokaryotic genomes[177]. Prokaryotic Agos are classified regarding the presence or not of the PAZ domain[190], differencing between long and short pAgos (Fig. 1.11). Long pAgos are comparable to eukaryotic Agos and may contain a complete or disrupted catalytic site, whereas short pAgos variants only comprise MID and PIWI domains and usually enclose an incomplete catalytic site[129]. DNA-specific predicted nucleases often cluster with pAgos harbouring an incomplete catalytic site, empowering the endonucleolytic activity[129]. In addition, genes coding for pAgo proteins co-exist and are correlated with genes coding for proteins involved in host defence processes, supporting further defence roles of pAgos[103,129].

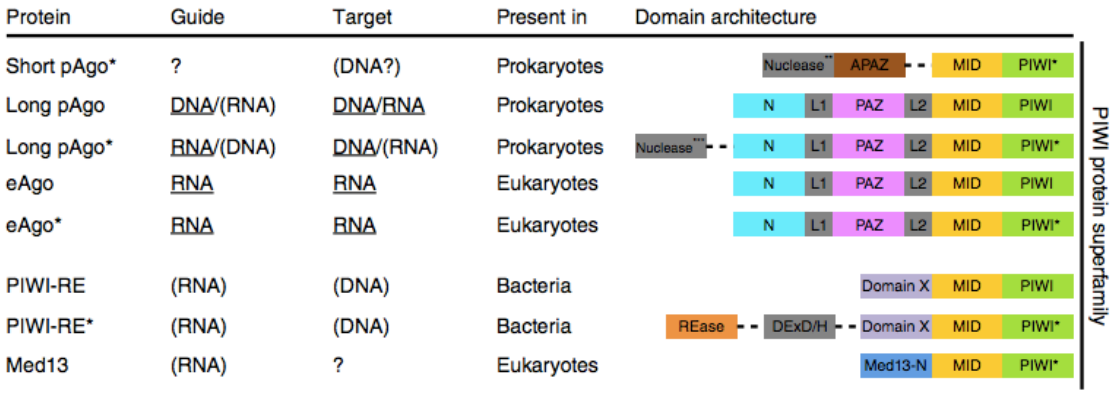


FIGURE 1.11. Diversification of Argonaute proteins based on PIWI domain architectures. As observed, despite all Agos interact with 5'-phiosphorylated oligonucleotide guides, some pAgos have higher affinity for DNA guides than for RNA guides. Ago proteins harbouring incomplete tetrad catalytic site are identified with an asterisk. Underlined data has been published whereas information within parentheses is foretold. Predicted DNA-specific nucleases from families Sir2, Mrr or TIR families are indicated by [a, b]. Adapted from Swarts *et al*[190].

In vitro studies have evidenced that, unlike eukaryotic Agos, some pAgos are able to bind ssDNA guides and use them to cleave RNA or/and DNA targets, diminishing DNA uptake yields during transformation[148,189]. This larger flexibility on

guide acquisition, target interference and catalytic mechanisms exhibited by pAgos, widens the relevance of pAgo's biological role, stimulating further research. Disentangling the molecular basis of further pAgo variants will undoubtedly contribute to gain insights on the evolution and potential biotechnological applications of pAgos[103,148,191,212].

1.3. THE *Thermus* GENUS

Thermus thermophilus belongs to the ancient *Deinoccoccus –Thermus* clade[150,210], ensued to be a profitable biological model in an assorted spectrum of biochemical, structural and evolutive studies. Besides, unique thermophilic traits of *Thermus* strains envisaged revolutionary breakthroughs which have certainly boosted their biotechnological applications[45].

The *Thermus* genus[26,25] is composed by ubiquitous thermophilic chemoorganitrophic bacteria, isolated from both natural (hydrothermal vents, hot springs, abyssal geothermal areas) and man-made (gold mines, hot water pipes, polluted streams) thermal environments[60]. *Thermus* bacteria are non-sporulating, bacillar, non-flagelated, filamentous at exponential growth and negative to Gram staining, although biochemical studies of its peptidoglycan place this genus as an evolutionary intermediate between Gram-negative and Gram-positive bacteria[56,96]. Actually, *Thermus* cell envelope is a multi-layered complex which embraces a slim peptidoglycan wall to which a secondary polymeric cell wall (SCWP) is covalently bound [39]. The cell is coated by an outer membrane which lays on a proteic S-layer, bestowing physical specific periplasmic compartments[39]. Occasionally, multicellular bodies may form by association of several cells sharing a single periplasmic space[26].

Several *Thermus* isolates grow aerobically to high yields in neutral to slightly basic complex media, at optimum temperatures from 60 to 75 ºC, although growth reports on some strains from *T. thermophilus* have fingered temperatures from 45 to 83 ºC[131]. Several strains within the genus can grow anaerobically, using nitrogen oxides as electron acceptors[138,157]. Many *Thermus* spp show an orange-yellowish pigmentation credited to the expression of membrane carotenoids which biological role may be linked to UV protection as a response to light[102]. Despite most strains prefer amino acids and carbohydrates as carbon and energy sources, many of them can grow in mineral media containing ammonia in addition to any sugar, pinpointing a truly complete biosynthetic metabolism[26]. *T. thermophilus* is the single halotolerant species of the genus[3]. Interestingly, *T. thermophilus* strains show a rare polyamine composition

which has been reported as essential for protein synthesis at high temperatures, as well as DNA and RNA stabilizers[152].

Currently, public sequences from five completed *T. thermophilus* genomes (strains HB27, HB8, JL-18, SG0.5JP17-16 and ATCC 33923) are available online. Besides, non-annotated contigs from various isolates are also available. Genome structure is highly preserved throughout *T. thermophilus* strains, entailing small genomes (less than 2.5 Mbp) composed by a well conserved chromosome (sized at 1.8–2 Mbp) and one or more megaplasmids, embracing the largest genomic flexibility and lower synteny between species. The extensive plasticity residing at the megaplasmids is probably owed to fluent lateral gene flow, suggesting that they may be acting as reservoirs for adaptation traits non-essential to life (e.g. carotenoid biosynthesis, defence systems, denitrification, etc.)[48]. Genomes are highly enriched in GC content (around 60-70 %), showing low gene duplications and a high degree of HGT-acquired genes of archaean origin, in addition to numerous insertion sequences and mobile genetic elements' related. These evidences may forecast the massive inter-domain DNA transfer, in addition to dynamic genomic rearrangements throughout evolution of these strains which could have rapidly conferred relevant adaptative advantages, like in other thermophiles[28,105].

Accessible work under laboratory conditions with some of the genome-sequenced *T. thermophilus* strains, specially HB27 and HB8, enhanced the development of a battery of genetic tools, unavailable for many other thermophiles[18,44,66,139]. High growth rates (doubling times of 40 min), the ability to form colonies on solid medium or its small genome which lacks on paralogs in other genomes are relevant features for this bacteria. Together, *T. thermophilus* role as cell factories for the overproduction of thermozymes and relevant protein complexes which can be further crystallized and, foremost, its extraordinary natural competence, support this organism as an ideal biological model.

1.4. HORIZONTAL GENE TRANSFER IN *Thermus spp*

Natural competence stands as a shared trait among several *Thermus* spp strains[113], able to uptake a broad range of DNA templates regardless of their sequence and origin[150]. Studies on *T. thermophilus* transformation leaded by Dr. Averhoff's group have fetched a detailed description of the competence machinery of the HB27 strain[29,81,165,168,174]. According to the latest update (Fig. 1.12), it is composed by at least 16 proteins, encoded in 7 different transcriptional units, distributed along the genome[16].

Interestingly, some of these proteins are more related to proteins of well-known conjugation models rather than in well known natural competence machineries[16]. In contrast to most studied models, this machinery is expressed in a constitutive fashion reaching transfers at extraordinarily high frequencies and uptake rates (40 kb·s^{-1} per cell)[174].

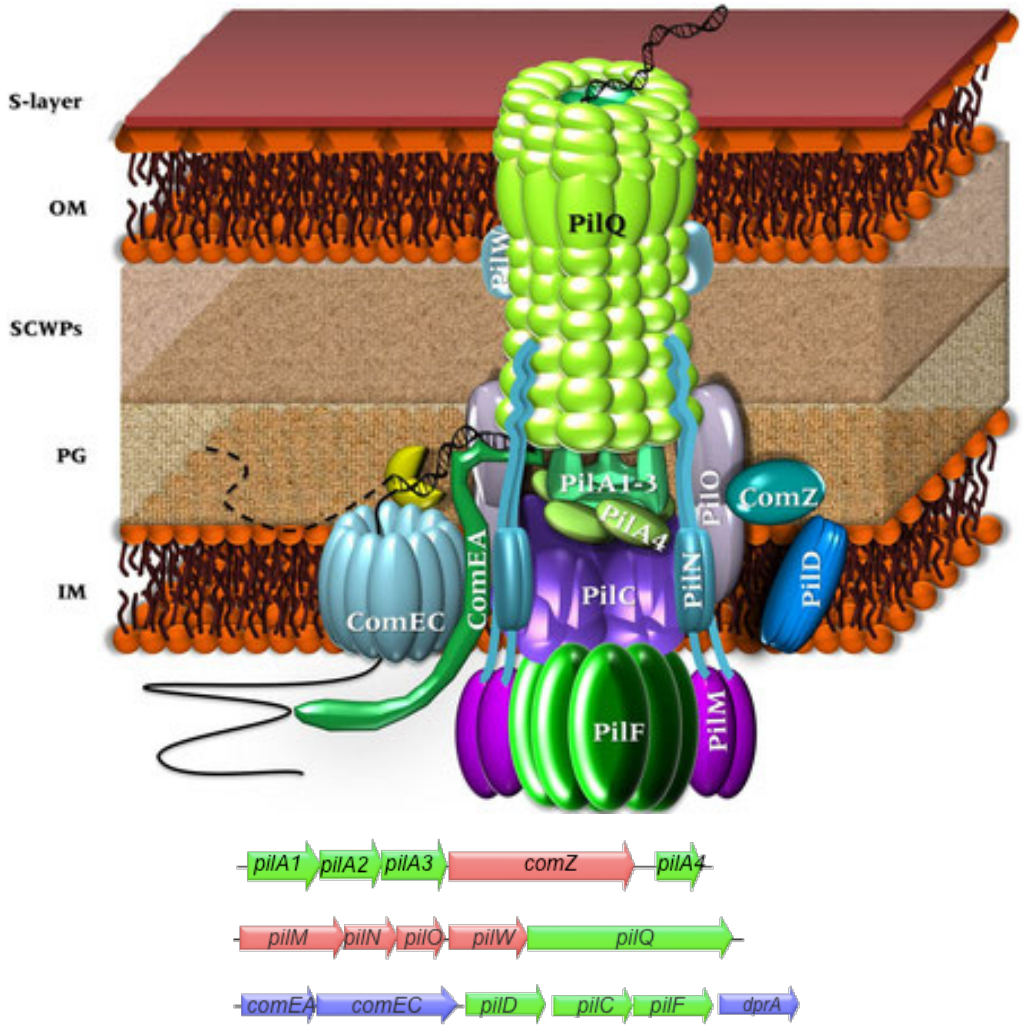


FIGURE 1.12. *T. thermophilus* translocator apparatus and genetic organization. DNA translocator model, as exposed in Dr. Averhoff's website. The traffic NTPase PilF powers the pilin polymerization assembling a fibre-like pilus at the periplasm, once the peptidase PilD has processed the pre-pilins A1-A4 into a mature state. Indeed, several reports suggest assembly of mature PilA4 into the pili extending from the inner membrane (IM) driven by PilF[168]. The pilin fibre crosses the cell wall from the peptidoglycan (PG) to the outer membrane (OM) via a dodecameric secretin ring formed by PilQ units, which localization is controlled by PilW, unique to the genus *Thermus*. Pilus retraction, commanded by PilF aided by PilT1/T2[168], facilitates DNA transport through the cell wall to the periplasm, where is collected by ComEA, which processes before delivering to ComEC at the IM, which transfers one strand of the DNA molecule to the cytoplasm, whereas the other is degraded by an unknown nuclease. DprA is involved in DNA processing at the cytoplasm. PilC, PilM, PilN and PilO are hypothesized to bind the multiprotein pilus assembly complex, coupling cytoplasmic and periplasmic parts of the T2SS and T4P systems. ComZ seems to be part of the transporter assembly scaffold[81,82].
Underneath, competence proteins assembling *T. thermophilus* transformation machinery are split into three groups, according to sequence and function homologies: in green, homologues to proteins from T4P systems; in blue, homologues to DNA-translocators-specific proteins from *Bacillus subtilis* and *Neisseria gonorrhoeae*; in red, homologues not related to any other known transformation system.

Hypothetical and structural reconstitution of the competence machinery has uncovered a complex DNA translocator apparatus, where the core is a type IV pseudopilus which spans through the whole cell by polymerization of pilins into long flexible fibres[16]. Oligomerization of the major dodecameric type II-like secretin PilQ forms a channel-pore like structure across which DNA could be transported[29]. Presence of pili is compulsory for an open state of the PilQ channel, which experiments major conformational shifts, accommodating to the extrusion of the pilus fibre within the pore through the outer membrane[88]. Besides, several proteins of the T4P play dual roles in piliation and competence, determining bacterial motility too[168].

Comparative analysis verified that this unconventional translocator apparatus is conserved throughout the genus, observing highest diversification among the pre-pilin operon[48]. In contrast, transformation efficiencies show drastic shifts between strains, from undetectable in *T. aquaticus Y51MC23* (Blesa, unpublished work) to the outstanding rates of 10^{-2} transformants/viable cell of HB27[113]. HB8 and PRQ25, which retrieved high scores of similarity to HB27 model, show a 2-log and 4-log decrease in competence frequencies compared to those obtained by the derivative $HB27^{EC}$, regardless of the DNA template[48].

Transduction has not yet been reported in *Thermus* spp, although more than a hundred *Thermus* bacteriophages were identified by Yu *et al* in 2006[214]. However, just few phages of *Thermus* spp have been subjected to any kind of study. Available information about *Thermus* bacteriophages is limited to ΦYS40, P23-77, P23-45, P74-26, IN93, Φ2119, ΦYB10 or W28P. Conversely, the conspicuous presence of multiple CRISPR systems in at least 5 strains of the *Thermus* genus, maximized in HB8 strain, which encodes 11 CRISPR loci classified in 4 different types of CRISPR systems[187]. This conspicuous presence of viral footprint may suggest that phage infection should be ordinary scenario upon *Thermus*, which would increase the odds of transduction incidence. Although yet to be described, transduction may play a relevant role in *Thermus* niche.

In the same line, a conjugation-like process among *T. thermophilus* cells was reported a decade ago, where chromosomal markers were transferred following a *Hfr*-like process at lower frequency rates in comparison to the plasmid-located denitrification genes[157]. Recently, both inter- and intra-strain DNase-resistant transfers between *T. thermophilus* cells have been reviewed, acknowledging HB27 strain as the universal receptor of such matings[48]. Further *in vivo* experiments showed that

conjugation-like process occurring in *T. thermophilus* was extraordinarily efficient, overtaking transfer frequencies achieved by natural competence. However, the molecular mechanisms underneath still remained uncovered as no homologues to proteins of classical conjugation systems could be identified in this strain in addition to those involved in natural competence. This suggested the existence of a novel unconventional conjugation apparatus[48]. Therefore, a major priority of this Thesis was to gain insights on the conjugosome operating in *T. thermophilus* cells, as well as the mechanisms and barriers shaping the conjugative events.

Chapter 2

Scope and objectives

CHAPTER 2: SCOPE AND OBJECTIVES

The onset of this PhD degree research relies on the knowledge that certain regions from *T. thermophilus'* genome can be transferred through a conjugation-like process. However, no information about the proteins involved, the factors affecting such transference, or the mechanism behind it, have been proposed yet. Thus, the main aim of this Thesis was to identify the major players and barriers involved in HGT in *T. thermophilus*. Therefore, the present dissertation is focused on meeting the following goals:

1. Explore lateral gene transfer routes in *T. thermophilus*, in particular, conjugation.

2. Decipher *T. thermophilus'* conjugosome following *in silico* and *in vivo* approaches.

3. Analyse the impact of insertion sequences in conjugative transfer and develop a transposon-based mutagenic tool.

4. Study the role of Argonaute in horizontal gene transfer in *T. thermophilus*.

Chapter 3

Materials and methods

CHAPTER 3: MATERIALS AND METHODS

3.1. MATERIALS

Bacterial strains as well as the newly generated ones are listed in table 3.1. Antisera employed are listed in table 3.2. Annex I endorses a list of the plasmids employed and constructed during this research. Oligonucleotides are recorded in Annex II. Buffers and reagents can be checked in Annex III.

TABLE 3.1. Bacterial strains employed in this dissertation.

Strain	Genotype	Phenotype/use	Reference/source
E. coli DH5α	*supE44 ΔlacU169 (Φ80 lacZΔM15) hsdR17, recA1, endA1, gyrA96, thi-1 relA1*	Ordinary cloning	Hanahan[97]
E. coli BL21 (DE3)	*F⁻ ompT gal dcm lon* HsdSB (r_B^- m_B^-) λ(DE3 [*lacI lacUV5-T7* gene 1 *ind1 sam7 nin5*])	Protein overexpression and purification	Rosenberg *et al*[162]
E. coli π	(F⁻) (lacJZYA) (lac) X74 *ΔuidA::pir116* recA1 *ΔphoA532 Δ(phnc?DEFGHIJKLMNOIP)33-30*	Transposase mutagenesis backbone plasmid	Demarre *et al*[69]
Thermus			
T. scotductus	*Thermus scotoductus* SA01	wt	E. Van Heeden
T. aquaticus	*Thermus aquaticus* DSM 625	wt	DSMZ
Taq *yfp*	*Thermus aquaticus* DSM 625 [pMKPnqosYFP]	Kan^R	This work
Tsco *yfp*	*Thermus scotoductus* SA01 [pMKPnqosYFP]	Kan^R	This work
Tsco *pyrE*	*Thermus scotoductus* SA01 [*TTC1380:kat*]	Kan^R	This work
T thermophilus			
HB8		wt	Y. Koyama
HB27 wt	*ATCC BAA-163/DSM7039*	wt	Y. Koyama
PRQ25		Complete denitrifer, wt	M. da Costa
SG0.5	*SG0.5JP17-16*	wt	US DOE JGI
NAR I	*[pTT27::nar]*	wt. NCE, partial denitrifer	Cava[43]
Fiji 3A1		wt	M. da Costa
VG7		Denitrifier	M. da Costa
$HB27^{EC}$	*HB27 ago::agoISTth7*	enhanced competence	This work

Δ*ago*	*HB27 Δago*	Argonaute less	Swarts *et al*[189]
HB27d	*HB27 [pTT27::nir-nor-nar]*	HB27 transformed with DNA from PRQ25 strain. Denitrifier	Alvarez *et al*[4]
27CM		HB27 Cm^R	This work
MD158	*HB27 [pTT27::*Δ(*TTP189-TTP079*)*]*	HB27 impaired in half pTT27 (from 186533 to 74904 is lost)	Carr *et al*[37]
ΔpilA4	$HB27^{EC}$ Δ*pilA4*	Non competent	This work
Δ*pulE*	$HB27^{EC}$ Δ*TTC1844*	*pulE* less. Non competent	This work
Δ*cptA*	$HB27^{EC}$ Δ*TTC1879*	Deficient in conjugative DNA donation	This work
Δ*nurA1*	$HB27^{EC}$ Δ*TTC1429*	Affected in viability	This work
Δ*cptB*	$HB27^{EC}$ Δ*TTC1430*	Impaired in conjugative DNA donation	This work
Δ*pilQ*	$HB27^{EC}$ Δ*pilQ*	Non competent	This work
Δ*herA*	$HB27^{EC}$ Δ*TTC0147*	Deficient in conjugative DNA donation	This work
Δ*repA2*	$HB27^{EC}$ Δ*TTP145*	Affected in viability	This work
SG0.5 *pyrE*	SG0.5JP17-16 [*TTC1380::kat*]	Kan^R. Denitrifier	This work
NAR I *pyrE*	NAR I [*TTC1380::kat*]	Kan^R. Denitrifier	This work
VG7 *pyrE*	VG7 [*TTC1380::kat*]	Kan^R. Denitrifier	This work
HB8 *pyrE*	HB8 [*TTC1380::kat*]	Kan^R	This work
27^{EC} *pyrE*	$HB27^{EC}$ [*TTC1380::kat*]	Kan^R	This work
Fiji *pyrE*	Fiji 3AI [*TTC1380::kat*]	Kan^R.Partial denitrifier	This work
C15	HB27 Δ*recA::kat [pTT27::* [Δ*TTP146::hyg] [pTT27::nar]*	Kan^R. Hyg^R. *RecA*$^-$. Denitrifier	This work
C23	HB27 Δ*recA::kat [pTT27::* [Δ*TTP146::hyg] [pTT27::nar]*	Kan^R. Hyg^R. *RecA*$^-$. Denitrifier	This work
C24	HB27 Δ*recA::kat [pTT27::* [Δ*TTP146::hyg] [pTT27::nar]*	Kan^R. Hyg^R. *RecA*$^-$. Denitrifier	This work
8CK1	HB8 [Δ*TTC1211::kat*]	Kan^R	This work
dCK1	HB27d [Δ*TTC1211::kat*]	Kan^R. Denitrifier	This work
CK1	HB27 Δ*TTC1211::kat*	Kan^R	This work
CK2	HB27 Δ*TTC1211::kat,* Δ*pilA4*	Kan^R. Non-competent	This work
CK3	NAR1 [Δ*TTC1211::kat*]	Kan^R. Denitrifier	This work
CK4	HB27 Δ*TTC1844::kat*, Δ*ago,* Δ*pilQ*	Kan^R. Non-competent	This work
CK5	$HB27^{EC}$ *TTC0638::kat*	Kan^R	This work

CK6	HB27^{EC} *TTC0893::kat*	Kan^{R}	This work
CK7	HB27^{EC} *TTC1415::kat*	Kan^{R}. Hyper-piliated	This work
CK8	HB27^{EC} *TTC1844::kat*	Kan^{R}. Affected in natural competence	This work
CK11	HB27 Δ*pilA4*::*kat,* Δ*ago*	Kan^{R}. Non-competent	This work
CK12	HB27 Δ*pilQ*::*kat,* Δ*ago*	Kan^{R}. Non-competent	This work
CK13	HB27 *pilF::kat*	Kan^{R}. Non-competent	Friedrich *et al*[80]
CK14	HB27 *pilA1-3::kat*	Kan^{R}. Non-competent	Friedrich *et al*[80]
CK15	HB27 *comEA::kat*	Kan^{R}. Non-competent	Friedrich *et al*[82]
CK16	HB27 *pilT::kat*	Kan^{R}. Non-competent	Friedrich *et al*[80]
CK17	HB27 *comZ::kat*	Kan^{R}. Non-competent	Friedrich *et al*[80]
CK18	HB27 *pilD::kat*	Kan^{R}. Non-competent	Friedrich *et al*[82]
CK19	HB27 *pilA3::kat*	Kan^{R}. Non-competent	Friedrich *et al*[82]
CK20	HB27 *ComEC::kat*	Kan^{R}. Non-competent	Friedrich *et al*[82]
CK21	HB27 pilT2::*kat*	Kan^{R}. Non-competent	Friedrich *et al* [80]
CK22	HB27 Δ*TTC1211::kat,* Δ*ago*	Kan^{R}	This work
CK23	HB27 Δ*TTC1211::kat*	Kan^{R}	This work
CK24	HB27 *TTC0638*::*kat,* Δ*ago*	Kan^{R}	This work
CK25	HB27 *TTC0638*::*kat*	Kan^{R}	This work
CK26	HB27 *TTC0858*::*kat*	Kan^{R}. Non-competent	This work
CK27	HB27 *TTC1621*::*kat,* Δ*ago*	Kan^{R}. Non-competent	This work
CK28	HB27 *TTC1844*::*kat,* Δ*ago*	Kan^{R}. Affected in natural competence	This work
CK29	HB27 *TTC1844*::*kat*	Kan^{R}. Affected in natural competence	This work
CK103	HB27^{EC} *TTC1450*::*kat*	Kan^{R}	This work
CK34	HB27*::ΔrecA::kat*	Kan^{R}. *recA*$^{-}$	This work
CK35	NAR1*::ΔrecA::kat*	Kan^{R}. *recA*$^{-}$. *Denitrifier*	This work
CK48	HB27^{EC} *TTC1430::kat;* Δ*pilA4*	Kan^{R}. Non-competent	This work
CK54	HB27^{EC} *TTC1879::kat;* Δ*pilA4*	Kan^{R}. Non-competent	This work
CK41	HB27^{EC} *TTC0638::kat,* Δ*pilA4*	Kan^{R}. Non-competent	This work

CK42	HB27^{EC} *TTC0893::kat, ΔpilA4*	Kan^{R}. Non-competent	This work
CK47	HB27^{EC} *TTC1429::kat; ΔpilA4*	Kan^{R}. Non-competent	This work
CK44	HB27^{EC} *TTC1844::kat; ΔpilA4*	Kan^{R}. Non-competent	This work
CK53	HB27^{EC} *TTC1839::kat; ΔpilA4*	Kan^{R}. Non-competent	This work
CK45	HB27^{EC} *TTC0147::kat; ΔpilA4*	Kan^{R}. Non-competent	This work
CK43	HB27^{EC} *TTC1415::kat; ΔpilA4*	Kan^{R}. Non-competent	This work
CK32	HB27^{EC} *TTC1429::kat*	Kan^{R}	This work
CK33	HB27^{EC} *TTC1430::kat*	Kan^{R}	This work
CK61	HB27^{EC} *TTC0147::kat*	Kan^{R}	This work
CK60	HB27^{EC} *TTC1879::kat*	Kan^{R}	This work
CK55	HB27^{EC} Δ*TTC1879::kat*	Kan^{R}	This work
CK56	HB27^{EC} Δ*TTC1879::kat, ΔpilA4*	Kan^{R}. Impaired in conjugation. Non-competent	This work
CK65	HB27^{EC} Δ*TTC0147::kat*	Kan^{R}	This work
CK66	HB27^{EC} Δ*TTC0147::kat, ΔpilA4*	Kan^{R}. Impaired in conjugation. Non-competent	This work
CK68	HB27^{EC} Δ*TTC1844::kat*	Kan^{R}. Affected in natural competence	This work
CK71	HB27^{EC} Δ*TTC1430::kat*	Kan^{R}. Affected in conjugation	This work
CK72	HB27^{EC} Δ*TTC1430::kat, ΔpilA4*	Kan^{R}. Non-competent. Affected in conjugation	This work
CK74	HB27^{EC} Δ*TTC1429::kat*	Kan^{R}.	This work
CK75	HB27^{EC} Δ*TTC1429::kat, ΔpilA4*	Kan^{R}. Non-competent	This work
CK59	NAR I [Δ*TTC1879::kat*]	Kan^{R}	This work
CK77	NAR I [*TTC0147::kat*]	Kan^{R}. Denitrifier	This work
CK79	HB27^{EC} Δ*TTC1844::kat, ΔpilA4*	Kan^{R}. Non-competent	This work
CK80	MD158 [Δ*TTC1211::kat*]	Kan^{R}. Lacks half pTT27	This work
CK81	HB27 Δ*TTC0656::kat*	Kan^{R}	This work
CK82	HB27 Δ*TTC0656::kat, Δago*	Kan^{R}. Affected in survival	This work
CK83	HB27^{EC} *TTC0656::kat*	Kan^{R}	This work
CK84	HB27^{EC} *TTC1026::kat*	Kan^{R}	This work

CK85	HB27^{EC} *TTC474::kat*	Kan^{R}	This work
CK86	HB27 *TTC1415::kat, ∆ago*	Kan^{R}	This work
CK90	HB27^{EC} *TTC0147::kat::sYFP, ∆pilA4*	Kan^{R}. Non-competent	This work
CK91	HB27^{EC} *TTC1879::kat::sYFP, ∆pilA4*	Kan^{R}. Non-competent	This work
CK92	HB27^{EC} *TTC1836::kat*	Kan^{R}	This work
CK93	HB27^{EC} *TTC1839::kat*	Kan^{R}	This work
CH1	HB27^{EC}*ΔTTC0313::hyg*	Hyg^{R}	This work
CH2	HB27 *pilQ::hyg*	Hyg^{R}. Non-competent	This work
CH3	NAR I *pilQ::hyg*	Hyg^{R}. Non-competent	This work
CH4	HB27 *∆pilA4::hyg*	Hyg^{R}. Non-competent	This work
CH5	HB27 Δ*TTC0313::hyg, ∆ago*	Hyg^{R}	This work
CH6	HB27 Δ*TTC0313::hyg*	Hyg^{R}	This work
CH14	HB27^{EC} *TTC1429*::*hyg*	Hyg^{R}	This work
CH15	HB27^{EC} *TTC1430*::*hyg*	Hyg^{R}. Affected in conjugation	This work
CH12	HB27^{EC} *TTC0147*::*hyg*	Hyg^{R}	This work
CH21	HB27^{EC} *TTC1879*::*hyg*	Hyg^{R}	This work
CpH24	HB27^{EC} [pMH::*TTC0147*::*sYFP]*	Hyg^{R}. Ectopic expression of HerA fused to sYFP	This work
CpH27	HB27^{EC} [pMH::*TTC1879*::*sYFP]*	Hyg^{R}. Ectopic expression of CptA fused to sYFP	This work
CH28	HB27^{EC} *TTC0147*::*hyg*::*sYFP*	Hyg^{R}. Single copy expression of HerA fused to sYFP	This work
CH29	HB27^{EC} *TTC1879*::*hyg*::*sYFP*	Hyg^{R}. Single copy expression of CptA fused to sYFP	This work
CH30	MD158 [Δ*TTC0313::hyg*]	Hyg^{R}. Lacks half pTT27	This work
CH31	HB27^{EC} *TTC0638::hyg*	Hyg^{R}	This work
CH32	HB27^{EC} *TTC0893::hyg*	Hyg^{R}	This work
CH34	HB27^{EC} *TTC0474::hyg*	Hyg^{R}	This work
CH35	HB27^{EC} *TTC0656*::*hyg*	Hyg^{R}	This work
CH36	HB27^{EC} *TTC1415::hyg*	Hyg^{R}	This work
CH37	HB27^{EC} *TTC1621::hyg*	Hyg^{R}. Non-competent	This work

CH38	HB27^{EC} *TTC1622::hyg*	Hyg^{R}. Non-competent	This work
CH39	HB27^{EC} *TTC1839::hyg*	Hyg^{R}	This work
CH40	HB27^{EC} *TTC1836::hyg*	Hyg^{R}	This work
CH41	HB27^{EC} *TTC1844::hyg*	Hyg^{R}	This work
PK1	HB27^{EC} [pTT27::*TTP046::kat*]	Kan^{R}	This work
PK2	HB27^{EC} [pTT27::*TTP085::kat*]	Kan^{R}	This work
PK3	HB27^{EC} [pTT27::*TTP140::kat*]	Kan^{R}	This work
PK4	HB27^{EC} [pTT27::*TTP167::kat*]	Kan^{R}	This work
PK5	HB27^{EC} [pTT27::*TTP191::kat*]	Kan^{R}	This work
PK6	HB27^{EC} [pTT27::*TTP219::kat*]	Kan^{R}	This work
PK7	HB27^{EC} [pTT27::*TTP211::kat*]	Kan^{R}	This work
PpK8	HB27^{EC} [pMKPnqosYFP]	Kan^{R}	This work
PK12	HB27^{EC} [pTT27::*TTP145::kat*]	Kan^{R}	This work
PK15	HB27^{EC} [pTT27::*TTP140::kat*], Δ*pilA4*	Kan^{R}. Non-competent	This work
PK17	HB27^{EC} [pTT27::*TTP191::kat*], Δ*pilA4*	Kan^{R}. Non-competent	This work
PK20	HB27^{EC} [pTT27::*TTP081::kat*], Δ*pilA4*	Kan^{R}. Non-competent	This work
PK19	HB27^{EC} [pTT27::*TTP208::kat*], Δ*pilA4*	Kan^{R}. Non-competent	This work
PK14	HB27^{EC} [pTT27::*TTP145::kat*], Δ*pilA4*	Kan^{R}. Non-competent	This work
PK26	HB27^{EC} [pTT27::Δ*TTP145::kat*]	Kan^{R}	This work
PK27	HB27^{EC} [pTT27::Δ*TTP145::kat*], Δ*pilA4*	Kan^{R}. Non-competent	This work
PK31	HB27^{EC} [pTT27::*TTP128::kat*]	Kan^{R}	This work
PK32	HB27^{EC} [pTT27::*TTP208::kat*]	Kan^{R}	This work
PK33	HB27^{EC} [pTT27::*TTP081::kat*]	Kan^{R}	This work
PK34	HB27^{EC} [pTT27::*TTP084::kat*]	Kan^{R}	This work
PH1	HB27^{EC} [pTT27::*TTP146::hyg*]	Hyg^{R}	This work
PpH2	HB27^{EC} [pMHPnqosGFP]	Hyg^{R}	This work
PH14	HB27^{EC} [pTT27::*TTP145::hyg*]	Hyg^{R}	This work
PpH15	HB27 [pMHPnqo::*TTP026*::sGFP]	Hyg^{R}. Ectopic expression of TtAgo fused to sGFP	This work

PpH16	HB27 [pMHPnqo::*TTP026*::sYFP]	Hyg^R. Ectopic expression of TtAgo fused to sYFP	This work
PH17	HB27 [pTT27::*TTP026*::sGFP]	Hyg^R. Single copy expression of TtAgo fused to sGFP	This work
PH18	HB27 [pTT27::*TTP026*::sYFP]	Hyg^R. Single copy expression of TtAgo fused to sYFP	This work
PH19	MD158 [pTT27::*TTP146::hyg*]	Hyg^R. Lacks half pTT27	This work
PH20	HB27d [Δ*nirS::hyg*]	Hyg^R. Partial denitrifier	This work
PH21	$HB27^{EC}$ [pTT27::*TTP128::hyg*]	Hyg^R	This work
PH22	$HB27^{EC}$ [pTT27::*TTP208::hyg*]	Hyg^R	This work
PH23	$HB27^{EC}$ [pTT27::*TTP191::hyg*]	Hyg^R	This work
PH24	$HB27^{EC}$ [pTT27::*TTP140::hyg*]	Hyg^R	This work
PH25	HB27 [pTT27::*TTP146::hyg*], Δ*ago*	Hyg^R	This work
PH26	HB27 [pTT27::*TTP146::hyg*]	Hyg^R	This work
PpH27	$HB27^{EC}$ [pMHPnqosYFP]	Hyg^R	This work
PH28	$HB27^{EC}$ [pTT27::*TTP081::hyg*]	Hyg^R	This work
PH29	$HB27^{EC}$ [pTT27::*TTP084::hyg*]	Hyg^R	This work
PH30	$HB27^{EC}$ [pTT27::*TTP085::hyg*]	Hyg^R	This work
CKH1	$HB27^{EC}$ Δ*recA::kat [pTT27::* [Δ*TTP146::hyg]*	Kan^R. Hyg^R. *recA*$^-$	This work
CKH2	HB27 Δ*TTC1211::kat [pTT27::* [Δ*TTP146::hyg]*	Kan^R. Hyg^R	This work
CKH3	$HB27^{EC}$ Δ*TTC1879::kat, TTC0147::hyg*	Kan^R. Hyg^R. Impaired in conjugation	This work
CKH4	$HB27^{EC}$ Δ*TTC1879::kat*, Δ*pilA4*, *TTC0147::hyg*	Kan^R. Hyg^R. Non-competent	This work
CKpH5	$HB27^{EC}$ Δ*TTC0147::kat*, Δ*pilA4,* [pMH::Pnqo::*TTC0147::hyg::sYFP*]	Kan^R. Hyg^R. Non-competent. Complementation	This work
CKpH6	$HB27^{EC}$ Δ*TTC1879::kat*, Δ*pilA4*, [pMH::Pnqo::*TTC1879*::hyg::sYFP]	Kan^R. Hyg^R. Non-competent. Complementation	This work
CKpH57	$HB27^{EC}$ Δ*TTC1879::kat,* Δ*pilA4*, [^{s}*TTC1879::hyg*]	Kan^R. Hyg^R. Non-competent. Complementation	This work
CKpH67	$HB27^{EC}$ Δ*TTC0147::kat,* Δ*pilA4*, [^{s}*TTC0147::hyg*]	Kan^R. Hyg^R. Non-competent. Complementation	This work
CKpH70	$HB27^{EC}$ Δ*TTC1844::kat,* Δ*pilA4*, [^{s}*TTC1844::hyg*]	Kan^R. Hyg^R. Non-competent. Complementation	This work
CKpH73	$HB27^{EC}$ Δ*TTC1430::kat,* Δ*pilA4*, [^{s}*TTC1430::hyg*]	Kan^R. Hyg^R. Non-competent.	This work

		Complementation	
CKpH76	HB27^{EC} Δ*TTC1429::kat,* Δ*pilA4*, [^{S}*TTC1429::hyg*]	Kan^{R}. Hyg^{R}. Non-competent. Complementation	This work
PKpH28	HB27^{EC} Δ*TTP0145::kat,* Δ*pilA4*, [^{S}*TTP0145::hyg*]	Kan^{R}. Hyg^{R}. Non-competent. Complementation	This work

3.2. MICROBIOLOGICAL METHODS

3.2.1. Bacterial culture conditions: growth and conservancy

Escherichia coli (*E. coli*) strains were grown at 37 ºC under shaking (180 rpm) in flasks or tubes filled up to 1/5 of total volume with LB medium [120] or on LB agar (1.5 % p/v) plates at 37 ºC for 24 h. For β-galactosidase-based recombinant selection plates amended with X-gal (40 μg·ml^{-1}) and IPTG (0.5 mM) were used.

Thermus strains were grown from 60 to 70 ºC under gentle shaking (150 rpm) in flasks filled up to 1/5 of total volume with TB medium[157]. Growth on plates was undergone in TB 1.5 % (p/v) agar, incubated for 48 h within wet chambers to prevent from desiccation.

Certain *Thermus* strains required special conditions for optimal growth. *T. aquaticus* DSM625 (Taq), *T. scotoductus* SA-01 (Tsco) and *T. thermophilus* (Tth) strains SG0.5 JP 17-16 (SG0.5) and Fidji A3-1 (Table 3.1) were grown in 1/2-1/4 diluted MilliQ water (MqH_2O) TB media, at 60 ºC under 150 rpm of rotational speed shaking.

Evaluation of the transfer of the denitrification island (chapter 5) required anaerobic cultivation. Anaerobic cultures were grown in 20-ml screw-cap tubes half-filled with MqH_2O TB media amended with potassium nitrate (40 mM), overlaid with mineral oil after inoculation.

In order to select resistant clones, media was supplemented with appropriate antibiotics at the following concentrations: Ampicillin (Amp; 100 μg·ml^{-1}), Kanamycin (Kan; 30 μg·ml^{-1}), Hygromycin B (Hyg; 100 μg·ml^{-1}), Bleomycin (Bleo; 3 μg·ml^{-1} for *E. coli* strains and 15 μg·ml^{-1} for *Thermus* strains), Chloramphenicol (Cm; 20 μg·ml^{-1}). *Thermus* M162 minimal medium was prepared as described by Degryse[68], micronutrients amended following Tanaka[192].

Growth of liquid cultures was monitored using a Hitachi U-2000 spectrophotometer at an optical density of 550 nm (OD_{550}) unless otherwise noted. For

bacterial growth curves or time-lapse curves, experiments were performed in triplicates with initial culture inoculated at OD_{550}= 0.05.

Short term strain conservation was performed on plates at 4 ºC for *E. coli* and at room temperature for *Thermus*. Longer storage was performed at -80 ºC in cryotubes supplemented with sterile glycerol 30 % (v/v). In addition, *Thermus* cells were kept at -20 ºC as pellets from 1 ml of stationary phase cultures.

3.2.2. Bacterial transformation

E. coli chemically competent cells were produced as described by Inoue[104] and performed following Hanahan's method[97]. For low transformation efficiency, DH5α cells were prepared following the $CaCl_2$ protocol[170], whereas highly competent cells of *E. coli* π strain were grown in SOB medium and prepared following the $RbCl_2$ protocol, employing TFBI and TFBII solutions (Annex III).

Thermus strains were transformed by natural competence[113] employing 0.01 - 1 µg of DNA which were added to 0.5 ml of exponentially growing culture (OD_{550} was approx. 0.35) in TB medium. Cells were incubated for 4 h at 60 or 70 ºC, and plated on selective TB plates, incubated for 48 h within wet chambers. Transformants were re-streaked on selective plates and further grown in liquid under each particular condition.

Competence defective *Thermus* strains were transformed by electroporation. Electro-competent cells were prepared from an overnight culture which was diluted 1:100 in TB and grown at 60 ºC with proper aeration until OD_{550} was 0.5. Cells were cooled down on ice and then gently centrifuged at 4 ºC. The pellet was washed twice with 1/10 of the culture volume of glycerol 10 % (v/v). Final re-suspension in glycerol 10 % (v/v) was aliquoted and stored at -80 ºC or directly used. 50 µl of electro-competent cells were incubated on ice for 30 min with 0.1 - 5 µg of DNA and then placed into a pre-chilled electroporation cuvette (0.2 cm thickness, Bio-Rad Gene Pulser® cuvette). Cells were subjected to 5 ms electric pulse under a 12500 V/cm electric field (Equibio, Easyject Plus D2000; 2500 V, 201 Ω, 25 F). Rapidly, 0.6 ml of pre-warmed media were added to the cells that were incubated in 12 ml tubes at the appropriate temperature and time before plating.

3.2.3. Transformation frequency assays

Transformability of different *Thermus* strains and derivatives obtained during this research was assayed implementing two variants of transformation tests. A qualitative approach, examining the ability to uptake free DNA was tested by culture spots. Briefly, 10 µl of an overgrown culture were laid on top of selective agar plates

and topped with 20-200 ng of naked DNA. Plates were incubated at 60-70 ºC for 48 h in wet chambers.

Quantitative transformation assays were carried out when required. For that, overnight cultures were re-inoculated and grown at 60-70 ºC to an OD_{550} of 0.35, when DNA (10-1000 ng) was added, as described above. Transformation frequencies were measured as the number of colony forming units (CFU) on selective plates per viable cells of the transformed strain. When relevant, transfer frequencies were related to the mass of DNA transformed. Experiments were repeated, 3 to 8 times to reach confident results.

3.2.4. Conjugation assays

The ability to transfer DNA *via* conjugation was assayed both qualitatively and quantitatively. For qualitative conjugation assays, 10^7 *Thermus* cells were laid on top of double selective agar plates and topped with 10^7 cells of the counterpart strain. One unit of DNase I (Roche) was supplemented to prevent putative transformation with DNA from lysed cells. In general, 1:1 ratio was employed unless specific tests were accomplished.

Quantitative approach included conjugation assays performed both in liquid and on solid plates. Most frequently, mating experiments involved overnight growth to saturation of *Thermus* strains. Then, 100 µl of each mating pair (unless other conditions were indicated) were mixed in the presence of 5 units of DNase I (Roche). Cells were centrifuged for 4 min at 5,000 rpm and resuspended in 10 µl of TB medium amended with 5 more units of DNase I and laid on sterile nitrocellulose filters (0.22 µm; Protran BA85, Whatman) which had been placed on pre-warmed TB agar plates. After 5 h at 60 ºC, filters were soaked in 1 ml of TB medium and vigorously vortexed to detach the cells. Appropriate serial dilutions were plated onto selective TB agar plates, which were then incubated at 60-70 ºC for 48 h inside wet chambers. Conjugation frequencies were referred as the ratio of the number of CFU resistant to both antibiotics to the number of CFU corresponding to one of the mates, routinely, the chromosome-labelled one.

Matings were also performed on liquid TB in presence of DNase I. Standard experiments involved 100 µl of each mating pair (unless otherwise indicated) which were mixed in 1.5-ml eppendorfs in presence of 5 units of DNase I, centrifuged for 4 min at 5,000 rpm and resuspended in 100 µl of TB medium, always supplemented with 5 units of DNase I and incubated for 5 h at 60 ºC under shaking (180 rpm). After incubation, appropriate dilutions were plated on selective plates, as described above.

In addition, variations of conditions of conjugation experiments involved assays using different ratios of mates at different growth stages, variations of the incubation time and temperature.

Conjugative transfer of the denitrification island was also performed adapting the ordinary transfer protocol to anoxic conditions in order to select for its transfer. Briefly, overnight cultures grown at 70 ºC were harvested and washed together in TB medium supplemented with 5 units of DNase I. 60 µl of the mix were laid onto 0.22 µm nitrocellulose filter and incubated at 60 ºC for 24 h. Cells were detached, resuspended in 1 ml of TB medium and added to 10-ml TB medium in anaerobiosis tubes, amended with potassium nitrate (40 mM) and the appropriate antibiotics for selection. Tubes were incubated for 48 h at 70 ºC without agitation. Individual colonies were streaked and test for denitrifying respiration and presence of denitrifying genes was confirmed by polymerase chain reaction (PCR).

3.2.5. Motility, adhesion and biofilm formation assays

Morphological characterization of derivatives generated in this study included the analysis of bacterial cell movement (or its absence) and their capacity to attach to surfaces.

Thermus motility is mediated by twitching[168], capacity which was qualitatively assessed by Coomassie staining of grown cells inoculated by injection on M162 minimal medium agar plates. Briefly, cells were grown for 72 h at 60-70 ºC on M162 agar plates which contained BSA[168]. After incubation, agar was stained with Coomassie blue (0.25 %) for 10 min which was later removed by careful washes with tap water. The colourless area radially depicted from the stab- inoculation point was the twiching zone, corresponding to the regions to where the cells had been able to move. Qualitative comparison of twiching zones between mutants as well as the diameter of this area is shown in table 6.2.

Qualitative cells' adhesion ability was checked by removing the agar from the plates and staining the Petri dish with Coomassie blue (0.25 %) for 10 - 15 min. Then, dye was removed and the plates were washed twice with tap water and twice with Tris-HCl (25 mM) and NaCl (100 mM). Cells bound to the plastic surface remained blue. Adherence diameters' for each strain tested are referred in table 6.2.

Biofilm formation assays were performed according to Salzer *et al*[169], based on the protocol of the microtiter plate assay by Koerdt[110] and O'toole[145]. Overnight cultures were diluted to and OD_{600} of 0.05. From each strain tested, an aliquot of 100 µl from

the diluted culture was placed on 96-well plate (flat bottom cell+, Sarstedt, Nümbrecht, Germany) which was then sealed with a Breathe-Easy membrane (Diversified Biotech, Boston) to prevent evaporation while providing good aeration, and incubated for 72 h at 60-70 ºC within a humid box. After incubation, the membrane was separated and OD_{600} was measured and registered. Medium with planktonic cells was carefully removed and the plate was washed twice with ~20 ml of MqH_2O. Staining of the attached cells was performed by adding 200 µl of crystal violet 0.1 % (w/v). The plates were incubated for 15 min at room temperature and then rinsed 3-4 times with water by submerging the plate in a tub of water and then dried for 2 h. A qualitative notion could be approached by observing whether there was biofilm formed and in that case where did the biofilm layer was localized. Quantitative assays involved solubilisation of the crystal violet by adding 200 µl of EtOH, followed by incubation for 30 min at room temperature. Finally, the absorbance was measured at 570 nm. Adherence was expressed as the ratio between absorbance at 570 nm and the OD_{600}.

3.2.6. DNA repair capacity tests

In order to test whether certain mutations had implied further impairments on the DNA repair machinery, ultraviolet (UV)-exposure tests (using a UV lamp UVG8T5 USA Sylvania USRAM StIII, Germany) were performed. 20 -40 µl of saturated cultures were laid onto TB agar plates in a continuous linear way. When adsorbed, and under sterile conditions, platted cultures were sequentially exposed to ultraviolet radiation (60 $J \cdot m^{-2}$) for 3 to 15 min, unprotecting every 30-60 s a small portion of the culture. Plates were incubated for 48 h at 65 ºC under darkness and damage was qualitatively estimated by the length of the linear grown culture.

In addition, quantitative estimates were gathered by plating serial dilutions of each strain, previously irradiated with 0 and 60 $J \cdot m^{-2}$ for 5 to 15 min in liquid at 15 cm distance to the UV source and further incubated at 65 ºC for 48 h. Ratios of CFU on irradiated plates over those non-irradiated were calculated.

3.2.7. Temperature-related growth assays

Each strain was grown overnight at 60 ºC. Cells were diluted in TB medium to OD_{550} of 0.02 and growth was monitored until each culture reached an OD_{550} of 0.8. Then, cells were incubated at 79.2 ºC and regularly, 100 µl-samples of the growing cultures were plated to check for viable cells. Final measure was taken after 24 h incubation at the highest temperature.

Additionally, the maximum temperature under which each culture is able to grow was tested on several *Thermus* derivatives. Each strain was over-day grown at 60 ºC, and re-inoculated for overnight growth at the selected temperature. Cells were diluted in TB medium to an starting OD_{550} of 0.02, since which growth was monitored at OD_{550} and CFU counting.

3.3. MOLECULAR METHODS

3.3.1. DNA techniques

In general, every protocol concerning DNA manipulation was based on those included in Sambrook[170].

Routinely, genomic DNA from saturated *Thermus* cultures was extracted and purified according to an adapted version for Gram negative bacteria with the DNeasy Blood & Tissue kit (Qiagen). However, for PCR checking, a faster method was used, based on the standard freeze-thaw technique[137]. In summary, 200 µl of saturated culture were centrifuged and resuspended in 150 µl of MqH_2O. Samples were subjected to 8 to 15 freeze (dried ice) - thaw (42 ºC) 3 min cycles, with vigorous vortexing after the thawing step. Finally, DNA was boiled for 5 min at 98 ºC and centrifuged to pellet the cell debris. Supernatants entailing the DNA were prepared for PCR analysis.

When high yields of genomic DNA were required, Marmur's protocol [133] was followed. Cells were centrifuged, washed in TES buffer (Annex III) and resuspended in Solution I, which was amended with RNase A (0.1 $mg \cdot ml^{-1}$; Roche) and lysozyme (0.1 $mg \cdot ml^{-1}$; Roche) in EDTA 0.25M pH 8 (Merck). After 1 h incubation at 37 ºC, 1 volume of TE buffer was added and cell lysis was driven at 56ºC for 1 h with N-Lauryl-sarcosine sodium salt (Sarkosyl, Sigma) in the presence of Pronase E (0.1 $mg \cdot ml^{-1}$; Merck). Subsequently, samples were incubated at 37 ºC for 30 min with proteinase K (3 $mg \cdot ml^{-1}$; Sigma). Phenol:chloroform:isoamylic (25:24:1) and chloroform: isoamylic (24:1) extraction was further used. DNA was precipitated with 2 volumes of pure EtOH (Merck) and $AcNH_4$ (100 mM; pH 5). DNA was resuspended in nuclease-free MqH_2O and checked by electrophoresis in agarose gel.

Plasmids were isolated from *E. coli* by the alkaline lysis method[170] or, with commercial kits.

DNA was amplified by PCR employing different polymerases, depending on the purpose. Routinely, *T. thermophilus* polymerase (Biotools B&M), Dream Taq DNA

polymerase (ThermoScientific) or NZYTaq polymerase (NzyTech) was employed. For high fidelity, Herculase II fusion polymerase (Agilent Technologies), *Pyrococcus furiosus* polymerase (Pfu, Biotools B&M) or Phusion® High-Fidelity DNA polymerase (New England Biolabs) were employed. Oligonucleotides used are listed in Annex II and PCR conditions were optimized following recommendations detailed by Ausubel *et al*[14]. Purification of DNA fragments, both from agarose gel excision or as PCR products, was performed with PCR-Prep kit (Promega) or High Pure PCR Product Purification Kit (Roche) following manufacturers' indications. DNA fragments were digested employing the appropriate restriction endonucleases, following manufacturers' indications (Roche, Fermentas or New England Biolabs). Cloning of certain DNA fragments required previous blunt-end conformation. In these cases, the DNA was treated with *Klenow* fragment from *E. coli* DNA polymerase I (Roche). Purified digested DNA fragments were ligated in a 10- µl final volume reaction using the T4 DNA ligase (Promega or Fermentas), following manufacturers' indications. Occasionally, DNA was dephosphorylated to prevent religation after endonuclease restriction, employing Antarctic Phosphatase (New England Biolabs). Seldom, ligation of DNA fragments was accomplished by a Gibson Assembly reaction (New England Biolabs), which allows multiple joining of DNA fragments independently from restriction sites. Ligation mixtures were used to transform the appropriate *E. coli* strain.

DNA concentration was measured in a Nanodrop ND-1000 (Thermo Scientific) spectrophotometer and checked by electrophoresis in agarose gels (0.7-1.5 %, low EEO, Conda Pronadisa) with 1X TAE buffer (Annex III) and stained with 1 $\mu g \cdot ml^{-1}$ ethidium bromide. Electrophoresis was carried out at 100 V.

Every construction employed in this study is referred in Annex I and has been sequenced at specialised companies.

3.3.2. Cell fractionation techniques

Isolation and purification of extracellular vesicles (EVs) was performed following described methods[108,159]. Briefly, exponentially cultures grown at 60 ºC were filtered through 0.45 µm nitrocellulose filters (Whatman PROTRAN BA85) and the cell-free fraction was centrifuged at 6,000 x *g* for 20 min at room temperature to eliminate residual cells and large cell fragments. Pellets carrying remnant cells and debris were discarded. Supernatants were collected, filtered through 0.22 µm sterile nitrocellulose filters and ultracentrifuged at 150,000 x *g* for 2:30 h at 4 ºC (TST41.14 KONTRON SW bucket rotor) in a Beckman Coulter Optima L100XP ultracentrifuge (Beckman-Coulter, Palo Alto, CA, USA) using polyallomer Beckman tubes. Then, pellets were washed

twice with HEPES buffer (50 mM; pH 7.5) and centrifuged at 60,000 x *g* for 30 min. Final pellet was resuspended in 50 µl of phosphate buffer (50 mM; pH 7.5). Subsamples were then prepared for DNA measurements, DNase I treatment, DNA extraction and electron microscopy analysis. Control samples from each of the supernatants collected during the purification process were analysed for the presence of living cells and extracellular DNA (eDNA).

Quantification of DNA from each fraction (supernatant and pellet) of the vesicle purification protocol was measured with a Nanodrop device before and after the DNase I treatment (1 h incubation at 37 ºC; Roche). Quantification of vesicles-associated proteins was performed with a Micro BCA protein assay reagent Kit (Thermo Scientific).

The release of eDNA was assayed by taking 1 to 2 ml samples at different times along the growth curve. The cells were eliminated by centrifugation and the supernatant filtered through 0.45 µm nitrocellulose filters. Then, part of the sample was treated with DNase I (5 units, Roche, for 1.5 h at 37 ºC) whereas other remained untreated, working as a control of total DNA. eDNA was recovered by $AcNH_4$–EtOH precipitation[170] and quantified. When needed, functional integrity of this eDNA was examined by transforming 1 µg into $HB27^{EC}$ strain.

In addition, EVs- associated eDNA was extracted from EV fractions by phenol-chloroform and further precipitated[133,170], EVs-associated proteins were precipitated with acetone[170]. Briefly, 500 µl of each sample were supplemented with 1 ml of cold acetone and overnight incubated at -20 ºC. Samples were then centrifuged at 15,000 x *g* for 15 min at 4 ºC. Supernatants were carefully removed and pellets resuspended in 10–20 µl MqH_20. Subsamples were checked in acrylamide gels.

3.3.3. Protein techniques

3.3.3.1. Preparation of cell protein extracts

To prepare whole-cell lysates, 10^6- $5*10^8$ cells from *Thermus* or *E. coli* cultures were gently centrifuged and resuspended in Laemmli lysis buffer (Annex III). Samples were then boiled for 10 min and centrifuged at 15,000 x *g* to eliminate insoluble cell walls. Protein analysis was carried out by SDS-PAGE in 8-12 % acrylamide gels or, when needed, gradient acrylamide gels were employed.

When required, preparation of soluble and insoluble fractions from each cell extract was accomplished. Cells from 250 µl of each culture were gently harvested by centrifugation, resuspended in 250–500 µl of TE or Tris-HCl 25 mM pH 7.5 buffer and subjected to sonication (3 pulses of 1 min, 0.8 power, 0.9 intensity, 0.5 s of cycle

frequency, in Labsonic U, B. Braun device). The soluble fraction was separated by double centrifugation at higher speed (15,000 x *g* for 15 min at 4 ºC), amended with Laemmli buffer, boiled and subjected to SDS-PAGE separation.

The insoluble fraction containing the cell envelopes was washed twice with cold phosphate buffer (50 or 25 mM; pH 7.5) and centrifuged at maximum speed following aforementioned conditions. Pellet was resuspended in SDS 1 % with EDTA 10 mM, boiled for 10 min and centrifuged to discard the peptidoglycan fraction. Supernatant was amended with Laemmli buffer supplemented with EDTA, and boiled again for 10 min.

A mix of known molecular weight markers were employed as standards (SDS-PAGE Molecular Weight Standards, Low Range, Bio-rad; Prosieve Color Protein Markers, Lonza). Protein visualization was achieved by staining with Coomassie blue.

3.3.3.2. Protein immunodetection (Western-Blot)

After SDS-PAGE electrophoretic separation, proteins were electro-transferred to PVDF membranes (Immobilon-P, Millipore) at 3 $mA \cdot cm^{-2}$ for 30 - 60 min in a Trans-blot SD Semi-dry Transfer Cell (Bio-Rad) in presence of transfer buffer (Annex III). Subsequent incubation in TBS-T containing 3 % skimmed milk (p/v) for 1 h at room temperature was followed by 1 h incubation with the primary antibody diluted in TBS-T at the stated dilution in each case. Sometimes, overnight incubation at 4 ºC was performed. Antisera employed in this research are enlisted in table 3.2.

PVDF membranes were thereafter washed three times with TBS-T for 5 min per wash to remove the excess of antibody. Later, secondary antibody incubation was performed for 1 h at room temperature in TBS-T employing the corresponding antibody conjugated to horseradish peroxidase. Normally, IgG goat anti-rabbit (GAR-HRP) was employed 1:5000 diluted in TBS-T. Membranes were washed again and incubated with developing solutions for chemiluminiscent detection. Then, autoradiographic film exposed for varying periods of time were developed using a KODAK X OMAT 2000 film processor.

In specific cases, immunoblotting was performed using nitrocellulose membranes to which the proteins separated by SDS-PAGE were electro-transferred employing the cathode and anode buffers (Annex III) for 1 h. Transfer was checked by staining the membranes with Ponceau S for 30 sec. Afterwards, membranes were blocked in 5 % skimmed milk diluted in TBS-T, analogous to PVDF membranes' treatment. Indeed, likewise proceedings were followed in relation to antibodies incubation and chemiluminiscent detection.

In certain occasions, multiple protein immunodetection was performed on the same membrane. After completing and developing the first Western-blot, antibodies' were detached by rotary incubation in stripping buffer (Annex III) at 50 ºC for 35 min. Subsequent intense washes with TBS-T enabled further implementation of Western-blotting of such membranes.

TABLE 3.2. Antisera employed in this study

Antisera	Characteristics	Reference
GAR-HRP	Goat anti-rabbit IgG conjugate from Bio-Rad. Horseradish peroxidase secondary antibody conjugate for chemiluminescent detection in Western-Blot	Bio-Rad
α-GFP	Polyclonal rabbit antibody employed for detection and localization of thermostable GFP reporter. Primary antibody in Western blot assays	Laboratory research
α-Kat	Polyclonal antibody employed for detection of thermostable *kat* cassette conferring kanamycin resistance. Primary antibody in Western blot assays	Laboratory research
α-TtAgo	Polyclonal rabbit antibody employed for detection and localization of *Thermus thermophilus* Argonaute (*TTP0026*) protein. Primary antibody in Western blot assays	Laboratory research
α-HerA	Polyclonal rabbit antibody employed for detection and localization of *Thermus thermophilus* HerA (*TTHA0522*) protein. Primary antibody in Western blot assays. Cross reactive with *TTC1879* and *TTC0147* from HB27 strain	This research
α-TtPilQ	Polyclonal rabbit antibody employed for detection and localization of *Thermus thermophilus* PilQ (*TTC1017*) protein. Primary antibody in Western blot assays	Bukhardt *et al*[29]
α-TtPIlF	Polyclonal rabbit antibody employed for detection and localization of Thermus thermophilus PilF (*TTC1622*) protein. Primary antibody in Western blot assays	Salzer *et al*[167]

3.3.3.3. Antibodies production

Most antisera employed in this research were available before this Thesis (Table 3.2). However, antisera against TtHerA and TtCptA were obtained by rabbit immunization with the purified protein encoded in *TTHA0522* locus from *T. thermophilus* HB8 strain, overexpressed in *E. coli* BL21 (DE3) from pET28b vector (Annex I). Protein was purified by affinity chromatography and used to immunize New Zealand rabbits, following normative proceedings (Directive 2010/63/EU). Controls of pre-immune antisera were tested and the titre of the antisera further determined. Antisera were stored aliquoted at -20 ºC.

3.3.3.4. Overexpression and protein purification of proteins

Protein overexpression and purification was conducted in *E. coli* BL21 (DE3) from pET28b (+) vector (Novagen), harbouring the *T. thermophilus* enzymes of interest. Cloning in this vector endorsed a 6 histidine extension tagging the proteins in

N- and/or C- terminus. Overexpression was performed in LB medium amended with Kan. The culture was diluted 1:100 in the same medium supplemented with 0.3 % glucose. Expression was induced by addition of IPTG (final concentration 1 mM) and incubated for 2 - 24 h at 37 ºC.

Proper induction of the protein of interest as well as its subcellular localization (soluble or insoluble fraction) was confirmed SDS-PAGE. Samples from a pre-induction and an induced culture were gently centrifuged (3,000 x *g* for 5 min), washed with TE 1X buffer (Annex III), resuspended in phosphate buffer (25 mM; pH 7.5) and sonicated. A 20 µl subsample of soluble and insoluble fractions were boiled for 10 min in Laemmli disruption buffer, centrifuged at high speed and supernatants were analysed by SDS-PAGE. Once induction was optimized and localization confirmed, protein purification was accomplished. Induced cultures were harvested by centrifugation, lysed by French press (137.9 MPa, Panda 2000 homogenizer, GEA Niro Soavi) and targeted proteins were purified by affinity chromatography employing columns with TALON CellThru Resin (ClonTech Laboratories, Inc.), following manufacturers' indications. After cell lysis, supernatants were heated at 70 ºC for 30 min in order to denaturalize thermolabile *E. coli* proteins, which would be then discarded by centrifugation (15,000 x g, 20 min, 4 ºC). Approximately, 2 ml of the resins were previously packaged, equilibrated and incubated with the samples for at least 2 h at 4 ºC in batch, so as to enhance interaction between Co^{2+} and the protein's his-tag. Columns were washed with 10 resin volumes of TALON washing buffer (Annex III) and then eluted with elution buffer (Annex III). Fractions with the purified protein were concentrated in centricons (Millipore) while buffer was changed by diafiltration to Tris-HCl (25 mM; pH 7.5). Stored aliquots were amended with 50 % glycerol at -20 ºC.

TALON purification protocol was adapted when the overexpressed protein of interest was not soluble and produced inclusion bodies (IB). Briefly, the pellet from a cell-extract was repeatedly washed with phosphate buffer (25 mM; pH 7.5) and solubilized with 2 ml of Triton 0.2 % (Annex III) for 30 min at 30 ºC under shaking. Detergent was then removed by several centrifugation washes and resuspended in 6M urea IB equilibration buffer (Annex III), for sample solubilization. Pre-equilibrated resin and cell-extract solution were incubated together overnight at 4ºC under gently mixing. Renaturalization was carefully performed by dialysis employing washing solutions entailing a degressive gradient of urea concentration, in order to guarantee correct refolding. 1 ml eluted samples were checked in SDS - PAGE gels and stored both at 4 ºC for immediate use, or in glycerol (50 %) to keep them at -20 ºC.

Protein concentration was measured using the Bio-Rad Protein Assay (Bio-Rad), following manufacturer's indications. In addition, proteins stained in SDS - PAGE gels and detected by immunoblotting were semi-quantified using the Quantity One software (Bio-Rad), following manufacturers' indications.

3.3.3.5. Proteomic analysis

Samples for further protein identification were prepared, amended with EDTA-Laemmli disruption buffer and heated at 90 ºC for 10 min. After boiling, samples were included in an in-gel sequencing grade trypsin digestion (Promega, Madison, WI), following manufacturer's indications. Gel extracted peptides were desalted and concentrated by OMIX C18 (Agilent Technologies), and embraced in subsequent mass spectrometry exploration. Results from the LC-MS/MS analysis were surveyed with SEQUEST search algorithm from Proteome Discoverer software (*v.* 1.4; Thermo Scientific) employing Uniprot's *T. thermophilus* HB27 database.

3.3.4. Molecular physiology techniques

3.3.4.1. Isolation of mutants of *T. thermophilus*

Plasmid construction approach implemented for mutant generation varied depending on the fate differencing single insertion mutants, where gene is disrupted but not removed; replacement mutants, where the gene(s) of interest is removed and replaced by an antibiotic resistance gene cassette; deletion mutants, where the sequence of interest has been removed from the genome; and fusion mutants, where the whole gene(s) of interest is in phase merged to a reporter or to an affinity tag.

Every construction was confirmed by PCR, digestion and sequencing. Mutants were also verified by PCR, immublotting and/or phenotypically. Further explanations are provided beneath (Fig. 3.1).

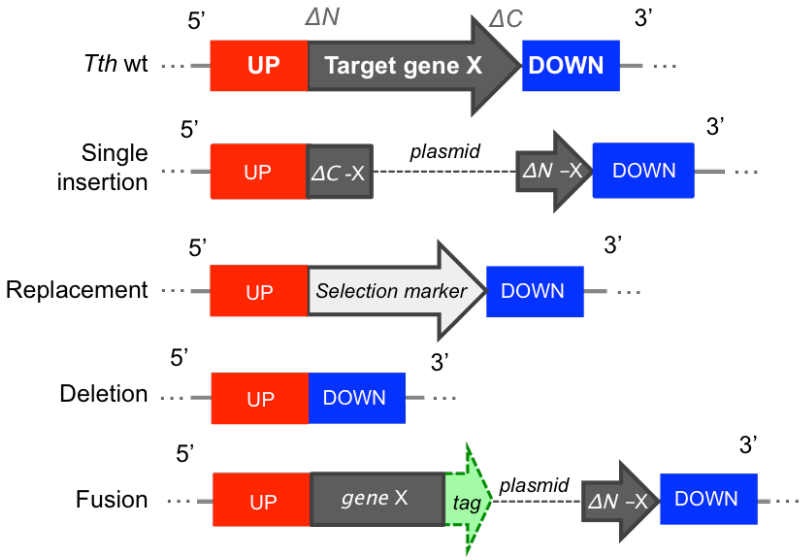


FIGURE 3.1. Genetic strategies for *T. thermophilus* mutants generation. Schematic representation of how insertion and recombination of each construct type has occurred in *T. thermophilus* genome. Routinely, selection marker was the *kat* or the *hph5* genes, conferring Kanamycin or Hygromycin resistance, respectively. The tag in fusion mutants was either a His-tag or a fluorescence reporter (sGFP, sYFP).

3.3.4.1.1. Single insertion mutants

Integrative vectors pK18 and pH118, selectable by Kan and Hyg, respectively, were employed as kick-off plasmids for mutant generation by homologous recombination. In general, a central part of each selected gene was amplified PCR (see Annex II for oligonucleotides) including flanking *EcoRI* and *HindIII* restriction sites for cloning into these vectors, unless otherwise specified. These constructions were directly transformed in *T. thermophilus* strains by natural competence or electroporation and transformants exhibiting Kan^R or Hyg^R after 48 h were checked by PCR.

3.3.4.1.2. Replacement mutants

Plasmid construction launched employing pUC19 and pUC18 vectors, non-replicative in *T. thermophilus*, in which upstream and downstream regions of the sequence coding for the gene to delete were cloned. Between these DNA fragments, the *kat* or *hph5* genes were cloned in downstream direction, conferring Kan^R or Hyg^R, respectively. Transformation of these constructs in *T. thermophilus* strains after linearization brought transformants which were further analysed by PCR to confirm the insertion of the construct and the absence of the target gene.

3.3.4.1.3. Markerless deletion mutants

Markerless “clean” deletion mutants were required for detailed physiological analysis of the most relevant mutations allowing as well the re-use of antibiotic markers. Kan^R replacement mutants were transformed with a pH118 plasmid harbouring the flanking regions of the deleted gene (Fig. 3.2). Proper integration of this plasmid by homologous recombination was confirmed by PCR, and transformants were grown in liquid TB medium under no selective pressure to favour plasmid loss by back recombination. Cultures were platted and colonies checked for antibiotic resistance. Those colonies which had lost both antibiotics resistances were PCR tested for the absence of the antibiotic marker and the target gene.

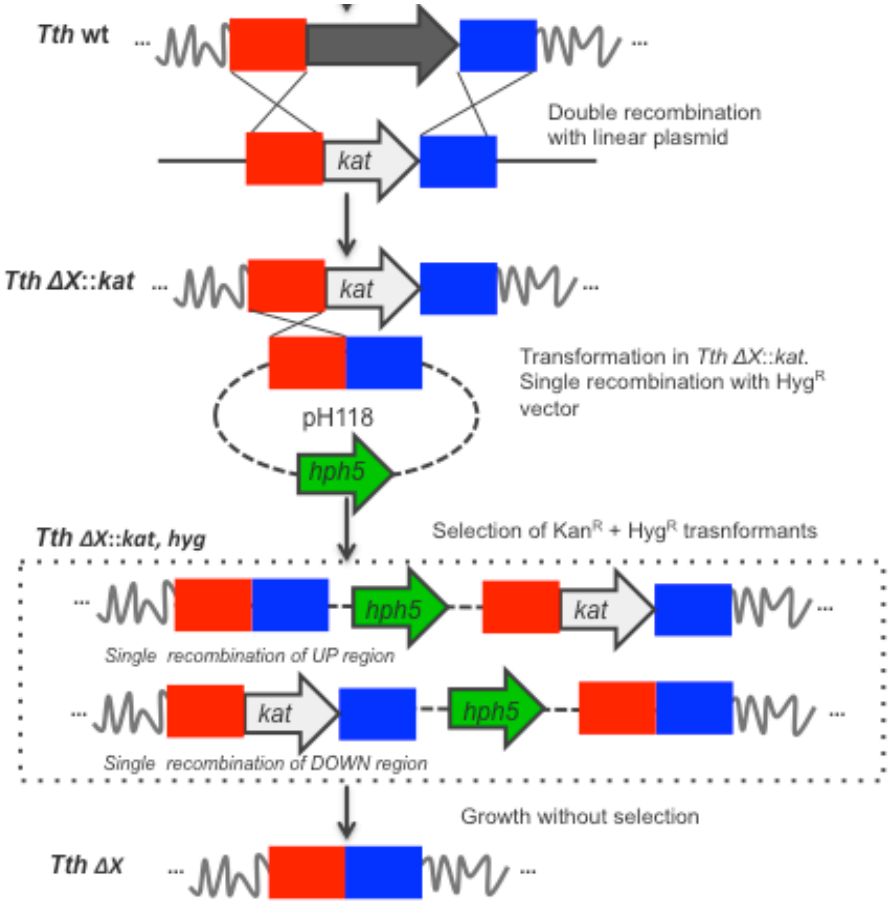


FIGURE 3.2. Markerless mutant generation by the “pop-in pop-out” strategy. Schematic representation of how markerless “clean” *T. thermophilus* mutants are generated by implementing a pop-in pop-out approach. First, a Kan^R replacement mutant is generated by double recombination of up and downstream regions of the region to delete. A second transformation is required. The Kan^R replacement mutant is transformed with a suicide Hyg^R vector harbouring the up and downstream regions aforementioned. Integration by single recombination, either by the up or downstream (represented inside the dotted box) conferred Hyg^R to the transformants. These Kan^R+Hyg^R transformants were subjected to several growth cycles without any selection pressure in order to stimulate markers withdrawal. Those which had lost both antibiotic resistances were PCR checked for proper deletion.

3.3.4.1.4. Fusion mutants

Localization and biochemical characterization of certain genes was explored by blending the C-terminal region of the protein of interest to a reporter, either a

fluorescent one, such as the thermostable sGFP or sYFP, or an affinity His-tag. Fusions were cloned in the suicide pK18 vector and used for single homologous recombination in *T. thermophilus*. Positive clones were checked by fluorescence in a Typhoon 9410 scanner (GE Healthcare) and further PCR and sequencing.

3.3.4.2. Mutant complementation assays

Verification of the phenotype produced by each mutation was accomplished by complementation with the wild type gene. For this, each selected gene was cloned into bi-functional plasmids pMH184/pMH185 and in pMK184, conferring Hyg^R (*hph5*)[141] or Kan^R (*kat*)[118], respectively. In these plasmids, the sequence cloned was constitutively expressed at different levels. Each of the mutants was transformed with one of these constructions, depending on previous markers selecting each strain, and the resulting double selective transformants were tested in HGT assays. Additionally, fusions of TtHerA, TtCptA and TtAgo to sGFP or sYFP were also included in complementation assays to test their functionality.

3.3.5. Enzyme assays

3.3.5.1. ATPase activity

ATP hydrolysis was estimated using the luciferin-luciferase ATP Bioluminescence Assay Kit CLS II (Roche), following manufacturers' indications. Different dilutions of purified enzymes in ATPase activity buffer (Annex III) were incubated with ATP (0.1 mM), MgSO4 (50 mM), NaCl (50 mM), for 1 h at 65 ºC and corresponding enzymatic activity was determined by the luciferin-luciferase coupled reaction at end-point assays, achieving 4 replicates per sample. Negative controls were performed without adding each substrate. Spontaneous conversion was also studied as negative control. The reaction analysed surveys the following principle:

$$ATP + D\text{-luciferin} + O_2 \rightarrow oxyluciferin + PP_i + AMP + CO_2 + light.$$

3.4. MICROSCOPY AND IMAGE ANALYSIS

3.4.1. Transmission Electron Microscopy

Different protocols ensuing this high-resolution microscopic technique were implemented for 2D and 3D protein modelling, the visual observation of the released EVs, and for the presence and distribution of pili structures along the cells of several *T. thermophilus* strains.

3.4.1.1. EVs observation

A 10-µl droplet sample of purified EVs from each strain studied were adsorbed onto ionized Collodion coated copper grids (300 mesh) and negatively stained with 2 % (p/v) aqueous uranyl acetate for 45 sec. The excess of solution was removed from grids, which were dried at room temperature. Grids were examined in a JEM 1010 transmission electron microscope (JEL, Japan) equipped with a TVIPS Tem Cam F416 (CCD SystemB) digital camera (Gauting, Germany) to take images using EM-Menu (*v* 4.09,53) software. Scale bars were included in each photograph.

3.4.1.2. Pili observation

Thermus cells grown on TB agar were carefully resuspended in 100 µl MqH_2O and adsorbed to copper grid (400 mesh), after twice washing steps with MqH_2O. Shadowing of the cells was carried out in a BAF 060 freeze-fracture system (BAL-TEC, Princ. Liechtenstein) with the following device specifications: pressure: $3–4*10^{-7}$ mbar; temperature: 28 °C specimen table temperature, angle: 25°, unidirectional; platinum/carbon thickness: 1.5 nm. Samples were examined in a EM 208S Transmission Electron Microscope (FEI, USA) at 80 kV. Image acquisition was executed with a Tietz TVIPS 1k x 1k slow-scan camera (Gauting, Germany).

3.4.1.3. Electron microscopy 3D reconstruction

Models for TtHerA and TtCptA were performed with the help and advise of Carlos P. Mata from J.R. Castón laboratory, at the CNB - CSIC centre (Madrid, Spain). Samples enclosing around 75 µg of the purified proteins, solved in filtered and degasified phosphate buffers, were incubated for 30 min at 65 ºC with 1-10 mM ATP and then prepared for negative staining, likewise the EVs preparations. Samples were visualized in a JEM 1010 transmission electron microscope, stabilized at 80kV. Images were recorded at a nominal magnification of 52000X using EM-Menu (*v* 4.09,53) software.

Digital processing of the samples was performed using the Xmipp software. Models were built from analysing more than 50,000 individual particles (86,506 for TtHerA and 103,550 for TtCptA) spotted in non-overlapping EM micrographs. Particles from 2X micrographs' subsamples (original size 2.44 $Å \cdot px^{-1}$) were automatically picked by the automatic routine application provided by Xmipp software followed by hand-cured supervision, extraction and normalization. Defocus ratio employed was calculated using the ctffind3 software, adjusting Fourier's space to accomplish the contrast transfer function correction (CTF) on the micrographs. 2D analysis was

performed using *CL2D* routine enclosed in Xmipp package. For 3D modelling, initial orientations and central part of the particles were determined using EMAN software[124] upon a noise model built with Xmipp. The model obtained was subjected to iterative angular refining performed following the projection matching routine enclosed in Xmipp, imposing hexameric symmetry and final resolution of each reconstruction was estimated adjusting the Fourier Shell correlation function to 0.5-0.3 (FSC= 0.5 or 0.3)[164].

3.4.2. Confocal microscopy

Experiments involving strains carrying sGFP and sYFP reporters were examined by confocal microscopy.

3.4.2.1. Sample preparation

Thermus cells were aerobically grown till DO_{550} was 0.2-1.3, depending on the particularities of each assay. When required, mid-exponential cells were incubated in presence of DNA or mixed with counterpart cells for 30 min to 1 h at 60 ºC. Generally, 10^6-10^7 cells were washed once with PBS 1X and fixed with 0.5-1.5 % paraformaldehyde (Merck) at 4 ºC. 50 µl samples were carefully laid onto slides previously covered with a ultra-thin 1 % (p/v) agar layer. Air-dried adsorbed cells were topped with cover slips carrying 50 µl Moviol mounting medium and preparations were dried prior to microscopic analysis. Duplicates for every sample were also prepared and kept in a cold light-protected device for 4 days maximum.

3.4.2.2. Digital image capture

Z-sections were taken employing immersion objectives 63X/1.4N.A for Zeiss LSM510 META coupled to an Axiovert 200M microscope and a 63X and 100X/1.3 for the Zeiss LSM710 vertical tied to a vertical AxioObserver microscope. Argon lasers ALEXA 488 (emission at 488 nm) and ALEXA 555 (emission at 555 nm) were employed to excite sYFP and sGFP labelled cells, respectively. Absorption spectrum ranged from 493 to 507 nm for sGFP and from 538 nm onwards for sYFP. Bright field snaps were also taken refined with Nomarski analyser. Chosen parameters fulfilled Nyquist criteria for image treatment. Final image design was completed with Image J software (Wayne Rasband, NIH, USA).

3.5. BIOINFORMATIC TOOLBOX

Sequence analysis, assembly, comparison and translation was performed using Vector NTI 10 (Invitrogen), Clone Manager (Sci-Ed software), Chromas Lite (Technelysium), Finch TV (Geospiza Inc.), 4Peaks (Mek & Tosj.com) and BioEdit (Thom Hall ibis Therapeutics). Search for homologous proteins was executed using BLAST and derivative tools at NCBI server and predicted functional interactions and network were checked at STRING v. 9.0 website. Sequence alignments were also performed using PRABI and CLUSTALW (EMBL-EBI European Bioinformatics Institute). Operons were inspected using the DOOR^2 database and OperonDB software. Restriction pattern inspection was verified using EnzymeX software. Sporadically, optimal primer design was tested at BMR genomics online portal.

Protein structures prediction, as well as calculation of their main features such as molecular weights and isoelectric points were accomplished using diverse tools enclosed in the ExPASy package[11]. In particular, subcellular location, transmembrane helix prediction and topology were performed implementing HMMTOP, TMHMM server 2.0, TMPred and TopPred 0.01, PSIPred and SOUSI. PSORTb *v.* 3.0.2 and CELLO *v.* 2.5 were employed in cellular localization research too. Three-dimensional protein prediction modelling was tailored using Phyre 2.0 online program and SWISS-Model based on retrieved homologous proteins from other organisms. Verification of such predicted structures was checked with RaptorX and I-TASSER. Additionally, protein structure was analysed using Pymol™ *v.* 1.3 (DeLano Scientiff LLC, 2009).

Phylogenetic trees as well as HGT predictions were performed using the online version of T-Rex software in addition to HGT-DB, IslandViewer, DOOR, Darkhorse and GC content information obtained at Gene Skew online portal.

Additionally, certain projects enclosed in this dissertation required implementation of specific software. For instance, putative ATPase identification was supported by pFAM and HMMer databases and potential ATPase involved in conjugation were submitted to multiple BLAST against CONJdb results. Within the project about the development of a transposon mutagenesis tool, synonymous sequences with optimal *T. thermophilus* codon usage adapted were obtained by displaying ELP program. Search for repeated sequences, putative transposases as well as identification of potential insertion sites was accomplished by implementing a battery of online databases linked to specific utilities such as ISfinder, Repeat Masker *v.* 3.2.6, Repseek, ORF Finder, TN-Pred, among others.

T. thermophilus complete genomes were browsed at BacMap server or at NCBI server otherwise. KEGG, Uniprot/Swiss-Prot databases were also surveyed.

3.6. STATISTICAL ANALYSIS

A set of statistical methods were used throughout the results of the present dissertation, all implemented by the statistical software SPSS ®Statistics v. 21.0 (SPSS Inc., Chicago, USA; 2008) and graphically represented employing SigmaPlot v.10 software. In all cases, results were considered significant if *p*–value < 0.05. Preliminary exploratory analysis employed inference tools, determining data distribution and outliers detection. Compliance of normality, homocedasticity and lineality was evaluated in order to select the adequate statistical tools. When any of these assumptions was not satisfied, non-parametric tests were assayed. To test whether conjugation and transformation transfer frequencies (calculated as aforementioned in the microbiological methods section) showed similar distributions, *Kruskal-Wallis* one–way analysis of variance tests were displayed. To examine differences among conjugation frequencies of transfers, student's *t*-test and Mann-Whitney-*U* tests were used. Differences among and within various loci HGT transfer frequencies were addressed by one-way ANOVA. *Wilcoxon*-tests were used when comparing sets of transfer frequencies of different strains to assess whether the transfer frequency mean differed. Simple linear regression test was implemented for modelling the relationship between loci and frequency of transfer. *Post-hoc* Tukey and Bonferroni tests were applied when convenient.

Chapter 4

Horizontal transfer of vesicle-protected eDNA among *Thermus* spp

Adapted from Blesa, A. and Berenguer, J. 2015. "Vesicle-protected extracellular DNA contributes to horizontal gene transfer in *Thermus* spp". *International Microbiology*. In press

CHAPTER 4: HORIZONTAL TRANSFER OF VESICLE-PROTECTED eDNA AMONG *Thermus* spp

4.1. SUMMARY

In this chapter, we demonstrate that culture supernatants of different strains of *Thermus* spp produce DNase-resistant extracellular DNA (eDNA) suitable for transformation of *T. thermophilus*. This eDNA is produced in in a growth-rate dependent manner, has a double-stranded nature, shows sizes similar to that of isolated genomic DNA and represents the whole genome of the producer strain. Protection against DNase is the consequence of the association of the eDNA to membrane vesicles which include a great diversity of cell envelope proteins with minor content of cytoplasmic proteins, suggesting for them a non-specific lytic origin. Access of the recipient cell to the protected eDNA depends on the natural competence apparatus. The low transfer frequencies observed indicates that this EVs-mediated eDNA transfer constitutes a minor contributor to HGT processes among *Thermus* spp. Nevertheless, their functional persistence over time could be evidencing their role as long-time and long-distance DNA vehicles among *Thermus* spp.

4.2. BACKGROUND

As shown in the introduction of this dissertation, some microorganisms exchange DNA through unusual mechanisms. This may be of particular relevance in thermal ambients, where tubular organelles and vesicle-like structures have been described to mediate eDNA exchange in a protected way[123,132]. Besides, the chances of occurrence of any of these alternative mechanisms in *Thermus* spp augmented when we detected DNase-resistant DNA transfer during HGT experiments (further details are provided in Chapter 5). Therefore, we looked for some of these unusual mechanisms regarded in the literature in cultures of *T. thermophilus*. Electron microscopy did not reveal any tubular structure beyond pili (Fig. 4.1) and no homologue to the described genes encoding nanopods or nanotubes was found encoded in its genome. Likewise, no evidence of gene transfer agents (GTAs) could be detected, which, in case of existence, it would be probably hampered by the scarce information about viruses predating *Thermus* spp. Finally, the presence of DNase-insensitive eDNA was examined by its association to vesicle-like structures. Indeed,

vesicle-protected eDNA has been reported as efficient DNA vehicles among several thermophiles[123,132]. Characterization of this eDNA and its horizontal transfer in *Thermus* spp is examined in this chapter.

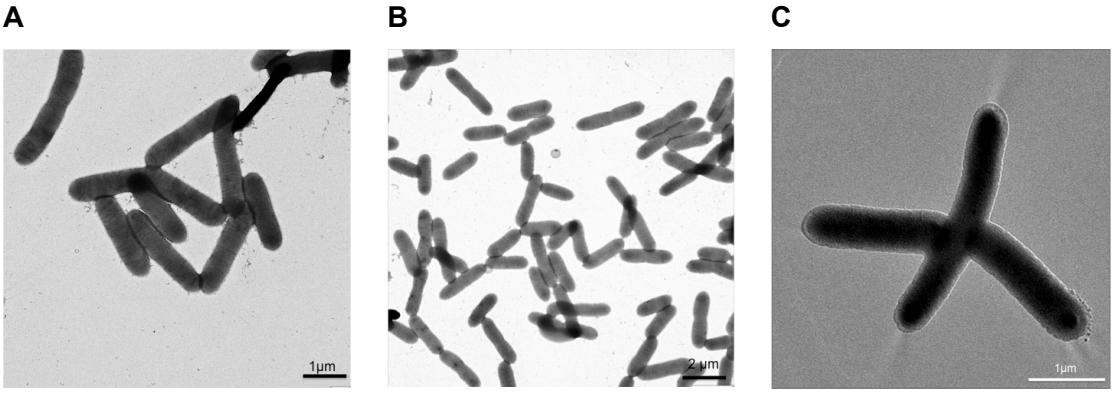


FIGURE 4.1. *T. thermophilus* has no apparent intercellular tubular structures. Electron micrographs of *T. thermophilus* HB27. **A.** HB27wt grown in rich medium at exponential phase. **B.** HB27 derivatives' mixed culture of CK1 and CH1 strains(see table 3.1 in Chapter 3), harbouring chromosomal kanamycin and hygromycin marker, respectively. **C.** *ΔpilQ* strain, impaired in transformation and deficient in pili structures.

4.3 RESULTS

4.3.1. Vesicle-protected extracellular DNA production in *Thermus* spp

In our first HGT assays, we observed low levels of transfer of genetic markers (around 10^{-6}-10^{-7} CFU/viable recipient) through a 0.22 µm nitrocellulose membrane in the presence of DNase I. This phenomenon evidenced the presence of DNase-resistant extracellular DNA (eDNA) in the media which could be taken up by the "recipient" culture. To provide numerical evidence of this observation, we treated with DNase I and filtered 1 ml samples of the growth media of three *T. thermophilus* HB27 derivatives, labelled with the *kat* gene cassette either in the pTT27 megaplasmid (PK7) or in the chromosome (CK1, CK11), and used them to transform the high transformation-efficiency *HB27 Δago* strain. As shown in figure 4.2, transformation was detected (on average 1.73 ± 0.68 $*10^{-6}$ CFU/viable recipient cell) with the growth medium from the three strains. Likewise, similar results were obtained when the DNase-treated supernatants of two HB27 derivatives impaired in a putative PulE-T2SS (CK8) or also in PilQ (CK4), the secretin from the competence apparatus, suggesting that DNase-protected eDNA was not secreted by these system.

We further analysed if production of transformable DNase-protected eDNA was a common trait within the *Thermus* genus. We examined culture supernatants from different strains of *Thermus* spp labelled with the *kat* reporter. To do so, we

transformed different *Thermus* strains with pAB153 plasmid (see Annex I). This suicide plasmid harbours a *kat* cassette inserted into the *pyrE* gene, which is highly conserved among the genus, leading to its integration by single recombination. As shown in figure 4.2.B, all *Thermus* spp (*pyrE::kat* derivatives) tested (*T. thermophilus* strains NARI, SG0.5JP17-16, VG7, and *T. scotoductus* SA1) produced enough DNase-resistant eDNA to allow its transfer by competence to HB27 *ago*$^{-}$ derivatives, although at very low rates (in average, 9.03 ± 2.12 *10^{-7}). No significant differences among strains could be detected (one-way ANOVA, *p*-value: 0.93), suggesting similar efficiency of transfer, thereby similar supernatants' eDNA content among all *Thermus* strains. We also tested the transfer of replicative plasmid pMKPnqosYFP from transformed strains. As shown in figure 4.2.C, this replicative plasmid was also transferred, although, again, at low rates (in average, 6.87 ± 2.03 *10^{-7}). These results demonstrated that both genomic and plasmidic DNA can be transferred through the media in a DNase-protected way.

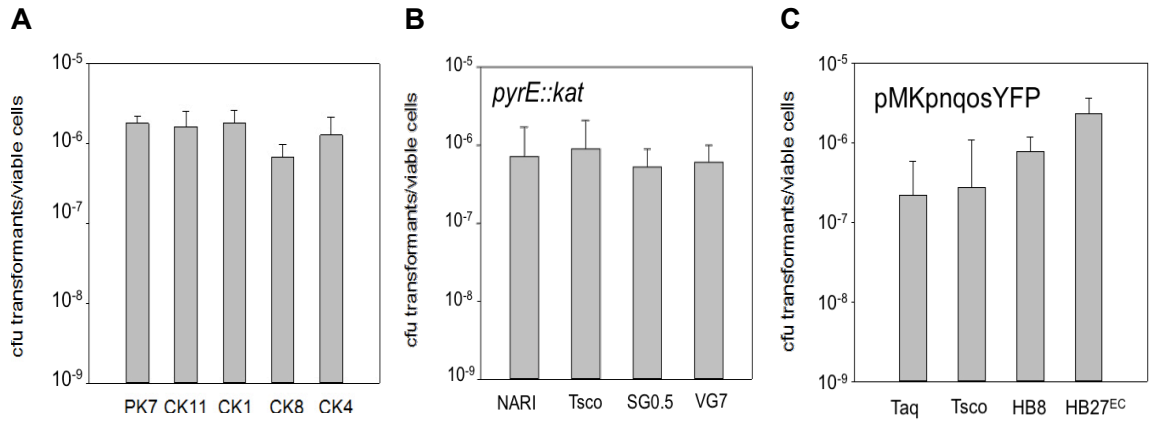


FIGURE 4.2. *Thermus* spp produces DNase-resistant eDNA. Transfer frequencies (expressed as the number of Kan^{R} colonies per viable cells) obtained in transformation assays involving 1 ml of filtered DNase-treated supernatants. Particularly, one ml samples of culture supernatants of each of overnight cultures of the donor strains were filtered and incubated with DNase I (5 units, Roche). Each sample was amended with extra 5 units DNase I and further used to transform 0.5 ml of growing cultures of *T. thermophilus HB27 Δ ago*. **A.** Transfer frequencies of two chromosome (CK1, CK11) and one megaplasmid (PK7) gene labelled with the *kat* cassette to the *Δ ago* strains. **B.** Transfer frequencies of supernatants from cultures of *Thermus* spp labelled with the *kat* cassette in the *pyrE* gene. Donor strains: *T. thermophilus* NARI (NARI), *T. scotoductus* SA01 (Tsco), *T. thermophilus* SG0.5JP17-16 (SG0.5), *T. thermophilus* VG7 (VG7). **C.** Transfer frequencies obtained of supernatants from different *Thermus* spp cultures harbouring plasmid pMKPnqosYFP. Donor strains: *T. aquaticus* YT1 (Taq), *T. scotoductus* SA01 (Tsco), *T. thermophilus* HB8 (HB8), *T. thermophilus* HB27^{EC} (HB27^{EC}). Error bars correspond to the mean standard deviation (n=3)

4.3.2. Characterization of DNase-resistant eDNA release

4.3.2.1. eDNA is produced along the cell growth

The production of eDNA was assayed by following the presence of eDNA in supernatants of four strains along its growth phases. The four strains selected were the wild type HB27 strain (HB27wt), its *Δago* derivative (*Δago*), and insertional mutants

CK8 (*pulE*-T2SS mutant) and CK11 (*ΔpilA4,* impaired in competence), both containing the *kat* cassette inserted into chromosomal genes. As shown in figure 4.3, production of eDNA could be detected along the whole growth curve in the four strains, being high in early exponential phase and decreasing with the growth rate. Actually, production of eDNA was exacerbated during growth in rich medium at 70 ºC compared to 60 ºC, whereas no eDNA was detected when the cells were grown in mineral medium at 60 or 70 ºC (Fig. S4.1, Annex IV). These data supported a relationship between fast growth and eDNA production, which ultimately insinuated that production of eDNA may be a consequence of cell lysis.

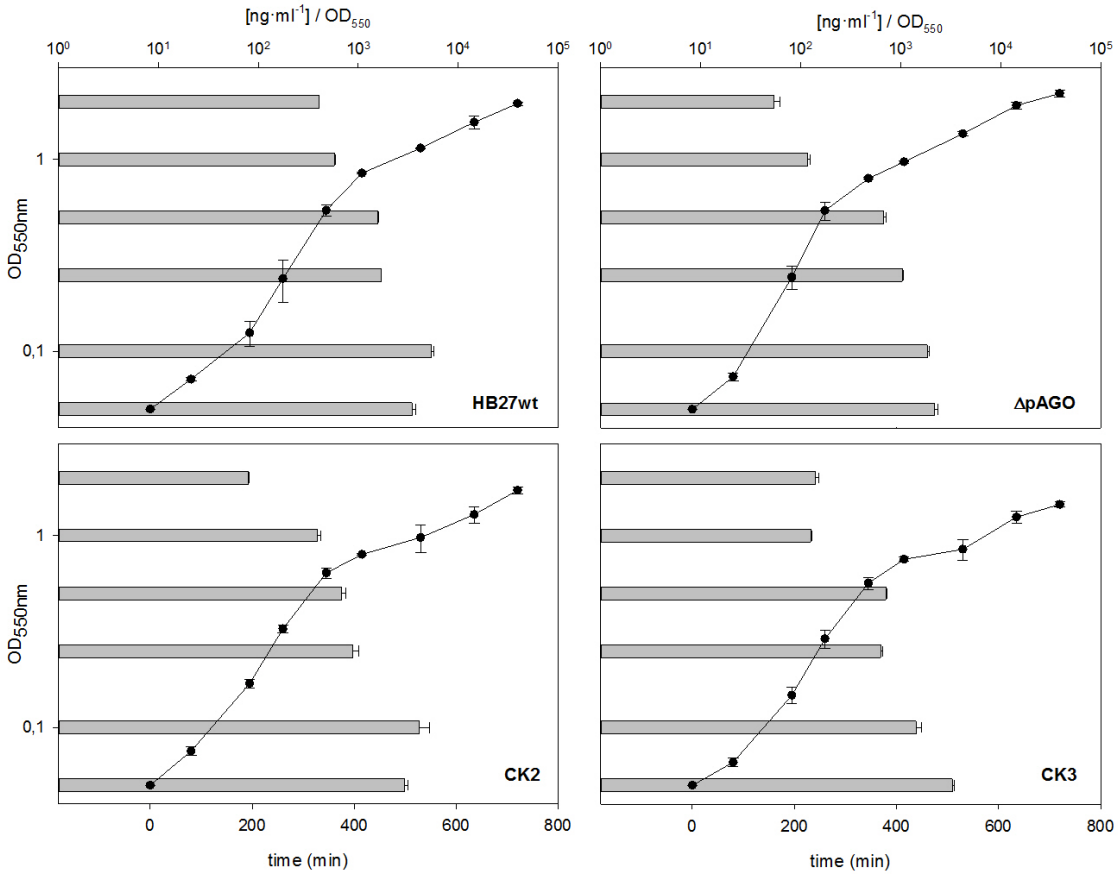


FIGURE 4.3. Production of eDNA is linked to the growth rate. eDNA released by the indicated strains of *T. thermophilus* along their growth in liquid TB rich medium was represented by bars as ng·ml^{-1} OD_{550nm}. Curves represented the amount of cells, measured as OD_{550nm}, along the growth cycle of each strain.

4.3.2.2. The eDNA produced by *Thermus* spp is representative of the whole genome

Data from figure 4.3 also suggested that eDNA production could be related to cell lysis during fast growth under rich medium. If so, it seemed likely that all regions of the cell genome should be present in the eDNA pool. Thus, to check whether *T. thermophilus* eDNA was limited to specific genome regions or, on the contrary, if it entailed heterogeneous DNA fragments, randomly scattered across the genome, we

analysed by gel electrophoresis the DNase-resistant eDNA from several growth media supernatants. As shown in figure 4.4, the eDNA from the HB27 strain and its CK8 and HB27^{EC} derivatives, had a size (around 20 kbp) similar to that of genomic DNA isolated by conventional methods, thus, suggesting that eDNA could represent the whole genome. To confirm this, we used two experimental strategies. On the one hand, we digested eDNA samples with the restriction enzyme *HindIII*, which has a relatively small number of cutting sites in the genome. If any defined fragment was detected it would suggest an enrichment in specific regions, whereas if products of undefined sizes were detected, this would point to an unspecific nature of the eDNA. The smear-digested pattern observed (Fig. 4.4.A) suggested that eDNA was representative of the whole genome. In parallel, we performed several PCR reactions on genes distributed along the chromosome (TTC genes) and the pTT27 megaplasmid (TTP genes) to confirm the random distribution of the whole genome within eDNA samples. In all cases we obtained positive results (Fig. 4.4.B). Thus, we concluded that eDNA included the whole genome of *T. thermophilus* in a random manner.

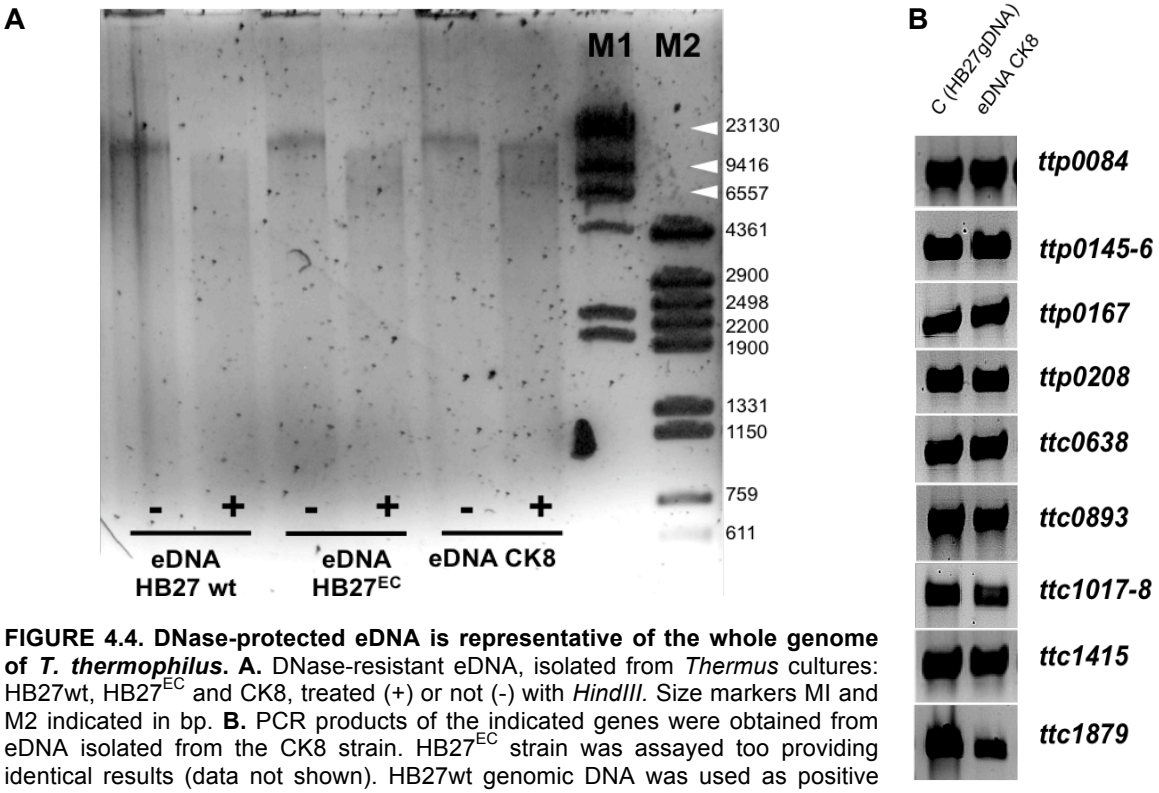


FIGURE 4.4. DNase-protected eDNA is representative of the whole genome of *T. thermophilus*. A. DNase-resistant eDNA, isolated from *Thermus* cultures: HB27wt, HB27^{EC} and CK8, treated (+) or not (-) with *HindIII*. Size markers MI and M2 indicated in bp. **B.** PCR products of the indicated genes were obtained from eDNA isolated from the CK8 strain. HB27^{EC} strain was assayed too providing identical results (data not shown). HB27wt genomic DNA was used as positive control. Oligonucleotides employed for PCR are described in Annex II.

4.3.3. Vesicle-like structures shelter eDNA

4.3.3.1. eDNA is protected within extracellular membrane vesicles

To investigate how eDNA was protected from DNase degradation we implemented cell fractionation protocols (see Chapter 3) upon the filtered supernatant

samples. Every fraction separated was incubated with DNase I and, afterwards, the presence of eDNA was quantified in a nanodrop device and checked in agarose gel (Table 4.1). Results indicated that most eDNA present in the culture supernatant was DNase sensitive. Indeed, 17.49 ± 6.35 % of the total eDNA precipitated from 1 ml of HB27wt culture supernatant was actually protected from DNase I degradation. Moreover, this DNase-protected eDNA could be sedimented by ultracentrifugation, but not by ordinary centrifugation, suggesting its possible association to small size vesicles structures.

TABLE 4.1. eDNA associate to vesicles is DNase-resistant and belongs to a sedimentable fraction by ultracentrifugation. Mean concentration (MEAN row) of eDNA ($ng \cdot \mu l^{-1}$) after filtering the supernatant (first column), present in the sedimentable fraction by ultracentrifugation (second column), the portion of DNase I resistant eDNA from this pellet (third column) and the eDNA extracted by phenol-chloroform from inside the isolated EVs (fourth column). Summarized data from HB27wt, CK8 and a mean value from the 6 strains tested, represented as the average and the standard error (MEAN±SD) and relative proportions of eDNA to the total (last row).

	Total DNA in supernatant	**eDNA associated to the pellet**	**DNase I resistant eDNA associated to the pellet**
HB27wt	461.30	85.11	68.80
CK8	426.53	86.33	67.23
MEAN ± SD [eDNA]	675.40 ±109.29	96.78 ± 20.46	95.68 ± 22.87
100-MEAN (%)	100	14.32	82.51

In order to check if the DNase-resistant eDNA sedimented by ultracentrifugation was associated to extracellular vesicles (EVs), we analysed the particulate fraction by electron microscopy (see Chapter 3). Results revealed the production of EVs in all the *Thermus* spp strains assayed. Figure 4.5 shows a representative image of such structures. In all cases, the sizes of such vesicles were rather heterogeneous. In general, despite their diversity, vesicles were almost spherical particles of ~30 to 80 nm in diameter composed of a dense core enclosed by a membranous layer, similar to those described in other Prokaryotes. In addition, as reported for *Thermococcales*[183], impurities and cell debris were present in multiple samples of vesicles.

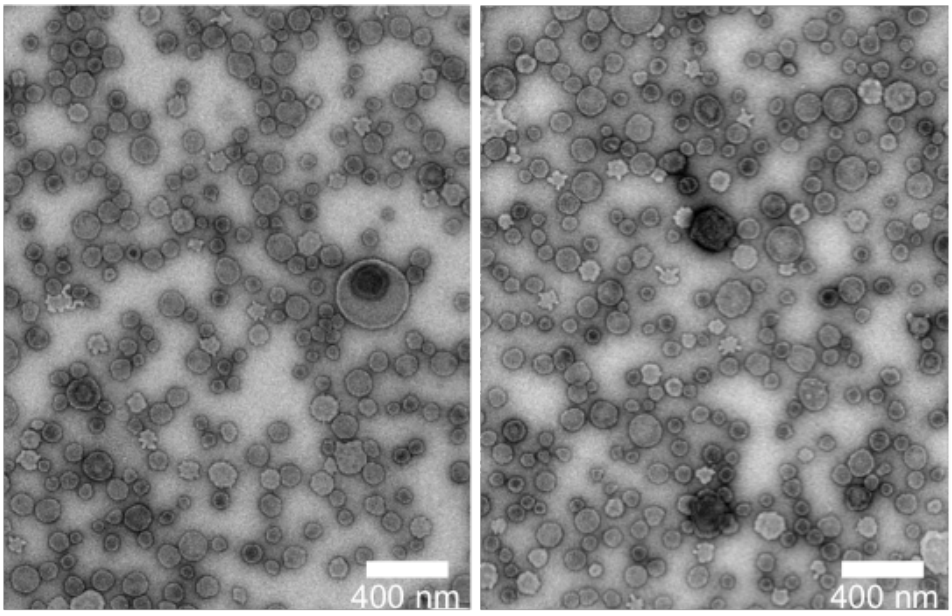


FIGURE 4.5. eDNA is protected within membrane vesicles. Representative TEM images of EV fractions of the HB27^{EC} and CK8 strains are presented. Samples adsorbed onto Collodion-coated copper grids were negatively stained with uranyl acetate. Samples were observed in a JEM 1010 transmission electron microscope. Bar represents 400 nm.

4.3.3.2. Protein composition of EVs

In order to analyse the composition of such EVs, a proteomic analysis was carried out with those isolated from the HB27wt strain. Mass spectrometry (LC MS/MS) of trypsin-generated peptides (see Chapter 3) identified 768 peptides related to 94 proteins (enclosed in the digital material). Then, we carried out a semi-quantitative analysis that gives a proxy of the protein content of the vesicles by using the frequency of peptides detected in relation to the corresponding molecular masses. We also investigated the proportion of peptides belonging to proteins encoded in the megaplasmid in respect to those in the chromosome (number of peptides divided by the bp of each molecule identified). We observed similar results among them as ratios were 4.1 and 5.2, respectively, supporting that proteins encoded for both molecules have similar chances to be found associated to the EVs. Besides, we bioinformatically predicted the subcellular location in specific cell compartments for each of the identified proteins. As shown in figure 4.6, most of the proteins for which a localization could be predicted belonged to cell envelope components (inner/outer membranes and periplasm). Among them, proteins addressed to the Omp family and S-layer peptides were predominant, serine proteases and ABC transporters were also very abundant, and competence proteins such as PilA4 and PilQ were identified too. However, a lower but significant fraction of proteins detected by mass spectrometry were identified as cytoplasmic components (14.2 %; Fig. 4.6). Regarding these data and the great diversity of proteins found, we concluded that these vesicles were most likely the

consequence of fortuitous cell lysis and not the product of a specific EVs-generation apparatus.

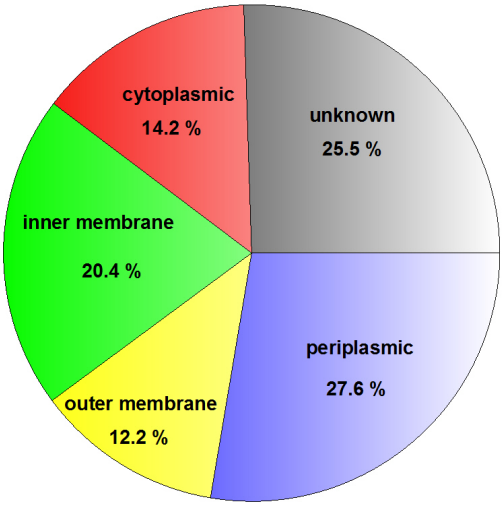


FIGURE 4.6. Subcellular localization of the proteomic composition of EVs. Relative abundance of the proteins associated to the EVs fraction, classified by their predicted subcellular localization, as predicted by P-SORTb software after LC MS/MS analysis.

4.3.3.3. eDNA is double stranded

In order to further characterize the eDNA embedded within the vesicles, we purified DNase I-resistant eDNA from the growth medium of exponential cultures of the HB27wt strain by chemical methods that destroyed the EVs and treated it with DNase or nuclease S1. As shown in figure 4.7, eDNA purified from supernatants or from EVs was sensitive to DNase (lanes 3 and 5) but not to S1 nuclease (lane 6). Control experiments with whole EVs showed the expected protection against both enzymes (lanes 2 and 4). Therefore, eDNA was essentially double stranded, also supporting the lytic origin hypothesis.

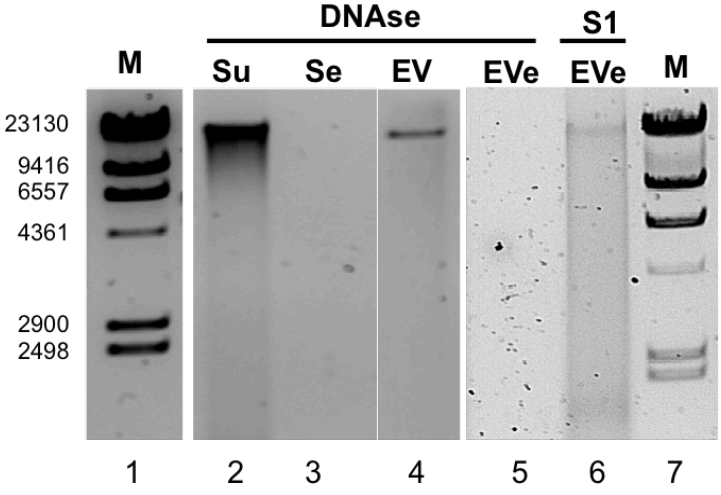


FIGURE 4.7. eDNA is double stranded and barrier-protected from nucleases. Samples were treated with DNase I (10 units, 1 h at 37 ºC) or Nuclease S1 (100 units, 1 h at 37 ºC) as indicated. Su) Supernatant from HB27; Se) eDNA extracted from Su by phenol-chloroform treatment and precipitated with ethanol; EV) Evs purified from Su by ultracentrifugation and washing; Eve) eDNA extracted from EVs; M) Molecular marker. size represented in bp.

4.3.4. Functionality in HGT of EVs-associated eDNA

We have already shown that eDNA was present in *Thermus* supernatants, protected from DNases action by its association to EVs and that these eDNA was susceptible to be horizontally transferred. We then questioned how this eDNA entered the cell. We initially performed a qualitative assay involving spot transformation of EVs isolated from the CK1 derivative to exponential cultures of HB27wt and competence mutant derivatives (*ΔpilQ, ΔpilA4).* Genomic DNA from CK1 was extracted and used as positive control. As shown in figure 4.8, *Thermus* strains impaired in transformation were not able to take EVs-associated eDNA. We also assayed transformation numerically, employing EVs and genomic DNA isolated from the CK1 and PK7 derivatives. Results exhibited low transformation rates with EVs compared to those obtained with genomic DNA. Interestingly, no differences could be observed between CK1 and PK7 EVs. These results were in agreement with results portrayed in figure 4.2, where no differences between transformation efficiencies could be observed between the different supernatants employed. All in all, these results confirmed that eDNA associated to EVs entered the cells through the natural competence apparatus.

Surprisingly, we repeated EVs transformation assays 20 months later, using the same EVs which had been kept stored at 4 ºC under sterile conditions. Yields obtained were lower but in an akin order of magnitude than those reported formerly, confirming that they were still functional in DNA transfer (data not shown). This finding supported the role of EV's as vehicles for HGT in *T. thermophilus* over long periods of time.

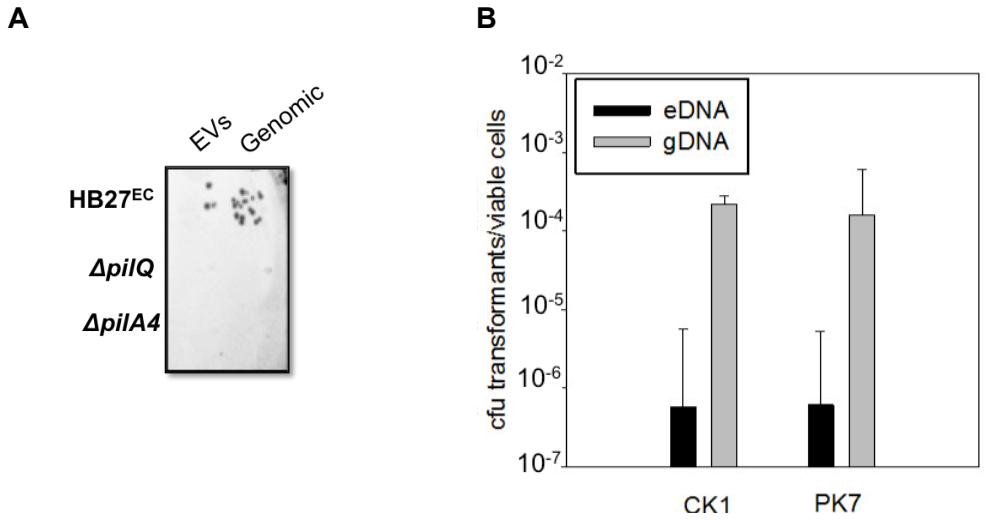


FIGURE 4.8. EVs-associated eDNA enters the cell through the natural competence apparatus. A. Spot transformation assay of HB27 *ago*⁻ strain and competence defective mutants *ΔpilQ* and *ΔpilA* (10 μl spots) with EVs fraction containing 300 ng of eDNA produced by cultures of the CK1 strain or with 20 ng of genomic DNA from the same strain in TBK plates. **B**. Transformation frequencies of *Δago* with EVs fraction (EV) from the CK1 and PK7 strains containing 500 ng of eDNA. Transformation controls with 20 ng of genomic DNA from same strains were carried out in parallel (gDNA). Transformation efficiency is expressed as the number of Kan^R CFU per viable cell. Error bars correspond to the mean standard deviation (n=3).

Chapter 5

Conjugation in *Thermus thermophilus*

Adapted from Blesa, A.; César, C. E.; Averhoff, B. and Berenguer, J. 2015. "Non-canonical cell-to-cell DNA transfer in *Thermus* spp. is insensitive to Argonaute-mediated interference" published in *Journal of Bacteriology. 197 (1), 138-146*

Parts of this chapter have been published in Álvarez, L.; Bricio, C.; Blesa, A.; Hidalgo, A. and Berenguer, J. 2014. "Transferable denitrification capability of *Thermus thermophilus*" published in *Applied and Environmental Microbiology. 80*(1), 19-28.

CHAPTER 5: CONJUGATION IN *Thermus thermophilus*

5.1. SUMMARY

In this chapter, we analyse an unconventional DNase I-resistant cell-to-cell DNA transfer mechanism which allows the exchange of genes at high efficiency between *T. thermophilus* strains. Despite substantial differences to standard conjugation-like processes, this cell-to-cell mechanism studied will be hereafter named as conjugation. We witnessed that several classical principles governing standard bacterial conjugation were broken in *Thermus* spp. The system fosters DNA transfers between isogenic strains, therefore being bidirectional, no apparent order of transfer could be deduced nor a clear *oriT* detected and no homologues to proteins involved in conventional conjugation could be identified. However, the mechanism has a strong preference (~10 fold) for genes associated to the megaplasmid compared to the chromosomal ones. The *T. thermophilus* conjugation system requires the components of the natural competence machinery in the "recipient" cell but not in the "donor". Based on this data, we propose a two-step "*push-pull*" DNA conjugation model where both mates display an active role. In this model, the donor actively pumps DNA through an apparatus independent from the transformation machinery. In the second step, the competence apparatus actively pulls the DNA into the recipient cell.

5.2. BACKGROUND

Pioneer demonstration of a conjugative process occurring among *T. thermophilus* was unravelled by interrupted-mating experiments[157]. In that study, the mobile nitrate conjugative element (NCE) was transferred in a DNase I-resistant fashion among non-isogenic *T. thermophilus* cells, following a process which resembled to an *Hfr*-like mechanism. Recently, Alvarez *et al*[5] confirmed the lateral transfer of the NCE as part of the nitrate denitrification island which included the NCE and the *nor-nir* cluster for consecutive nitrite and nitric oxide reduction[22]. This entire denitrification island spans a ~35 kbp region within a variable sector of the pTT27 megaplasmid, crowded of insertion sequences. Indeed, requirement of a highly processive and DNase-protected transfer process was compulsory in order to achieve a complete transfer of such large DNA fragments in a single step[5].

More recently, César *et al*[48] developed a battery of mating experiments which suggested that this DNase-resistant transfer was a common trait of the *Thermus*

genus, as these plasmid-associated denitrification islands were transferable among several *Thermus* spp. Besides, they demonstrated that transfer was not restricted to the mobilisation of the denitrification island[48], but was expandable to virtually any region within the HB27 genome, as long as there was intimate cell contact between two living cells. However, in HB27 one of the most efficient strains in conjugation[48], *in silico* analysis could not identify any homologues to components found in classical conjugosomes featuring Proteobacteria such as T4SS, beyond those already known to be involved in natural competence[81].

All in all, the proccesive DNase-resistant transfer of DNA among *Thermus* strains of any DNA genomic fragment in the absence of homologues to proteins involved in conventional conjugation systems suggested the existence of a novel class of conjugation mechanism among *Thermus* spp.

5.3. RESULTS

5.3.1. Chromosomal and megaplasmid-encoded genes are transferable by conjugation

In order to quantify and analyse in detail conjugation in *T.thermophilus*, we designed mating experiments involving the transfer of thermostable antibiotic resistance markers in a battery of mutants where the genes conferring resistance to Kanamycin (K) or Hygromycin (H) were inserted at different positions within the genome, either in the chromosome (named as CK or CH mutants) or in the megaplasmid (PK or PH mutant strains).

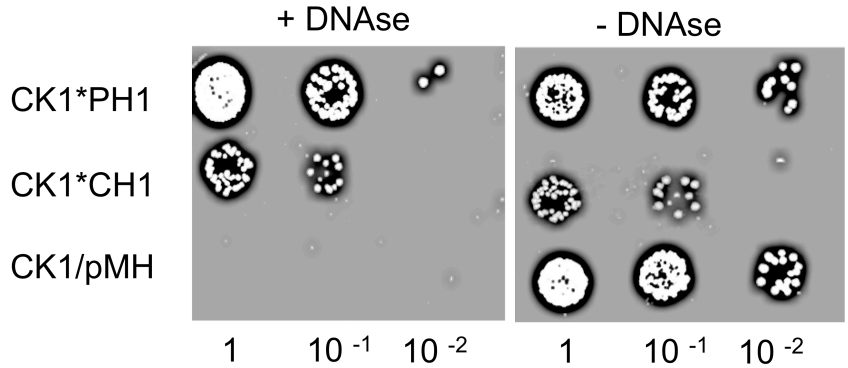


FIGURE 5.1. Conjugation in *T. thermophilus*. 10µl spots on double selective TBHK plates of serially diluted mating pairs between CK1 and a megaplasmid (PH1, first row) or a chromosomal (CH1, second row) Hyg^R strains, mixed in a 1:1 ratio. Bottom row shows serial dilution of Hyg^R+Kan^R CFU arisen from transforming CK1 with 200 ng of plasmid pMHPnqosGFP. Left panel, matings and transformation experiments were supplemented with 5 units of DNase I (Roche) before mixing.

As it can be observed in figure 5.1, DNA exchange between CK1 with PH1 and CH1 strains was insensitive to DNase I, whereas this enzyme abolished the

transformation of CK1 with plasmid pMH118. Therefore, both chromosomal and megaplasmid associated genes could be exchanged in a DNase insensitive way. Moreover, it was also noteworthy that conjugation involving a megaplasmid marker (CK1*PH1) was as efficient as transformation of the CK1 strain (CK1+pMH118), whereas conjugation involving chromosomal genes (CK1*CH1) was apparently less efficient (*i.e.*, lower transconjugants frequency rates).

Further characterization of the conjugation between CK1 and CH1 was addressed by varying several settings of the protocol (Chapter 3), including changes in DNase I concentration, time of DNase I addition, procedure performance (in liquid, in spot, cell mixes laid onto nitrocellulose membranes), etc. In all cases, frequencies were high and did not significantly vary, indicating that the conjugative transfer was robust and insensitive to DNase I (Fig. S5.1; Annex IV).

On the other hand, this experiment also confirmed that conjugation could occur between two isogenic cells such as CK1 and CH1. Indeed, high frequencies of DNA transfer were registered also in matings involving HB27wt mates harbouring small size plasmids such as pMHPnqosGFP and pMKPnqosGFP, respectively ($8.43 \pm 3.50 * 10^{-5}$).

In order to compare the efficiencies of conjugation and transformation, we performed matings between CK1 and CH1 in parallel to transformation of CK1 with a comparable amount of genomic DNA extracted from CH1 strain. Results shown in figure 5.2 evidenced that, with the same DNA template subjected to transfer (*TTC0313* chromosomal gene is replaced by the *hph5* cassette, conferring Hyg^R), the efficiency of conjugation was higher than that achieved by natural competence (*p*-value< 0.01).

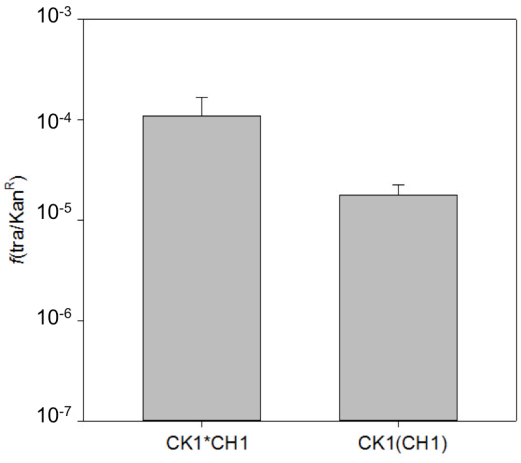


FIGURE 5.2. Conjugative transfer efficiency is higher than that exhibited by natural competence. Transfer frequencies were obtained after mating equal cell amounts of CK1 and CH1 supplemented with 5 units of DNAse I (Roche) [CK1*CH1]. In parallel, transformation of CK1 with 30 ng of genomic DNA isolated from strain CH1 [CK1(CH1)]. Frequencies represent an average of 9 independent experiments. Error bars correspond to the mean standard deviation.

To study if both transformation and conjugation processes could co-occur simultaneously, we mated CK1 and CH1 in nitrocellulose filters (Chapter 3) and supplemented the cell mixture with 1 µg of $Bleo^R$ plasmid pWUR. Although the actual

simultaneity could not be verified, growth of colonies in plates with the three antibiotics (Kan, Hyg, Bleo) suggested the co-occurrence of both processes although at low frequencies (13.34 ± 3.73 $*10^{-8}$). In contrast, frequencies estimated for conjugation (8.79 ± 3.31 $*10^{-5}$) or for transformation (16.50 ± 3.30 $*10^{-6}$) showed similar trends than those reported above, thus, suggesting no distortion among neither HGT processes.

5.3.2. Conjugation in *Thermus* mediates a preferential mobilisation of pTT27 megaplasmid associated genes

Results observed in figure 5.1 suggested a higher efficiency of transfer when pTT27 genes were involved. In order to further investigate this, we compared matings of CK1 with the chromosomally labelled CH1 strain, and the pTT27-labelled strain PH1. As shown in figure 5.3, CK1*CH1 transfer frequencies (6.1 ± 1.10*10^{-5}) were statistical significantly lower (~15 fold; *p*-value< 0.01) than those observed when involving a megaplasmid-labelled marker.

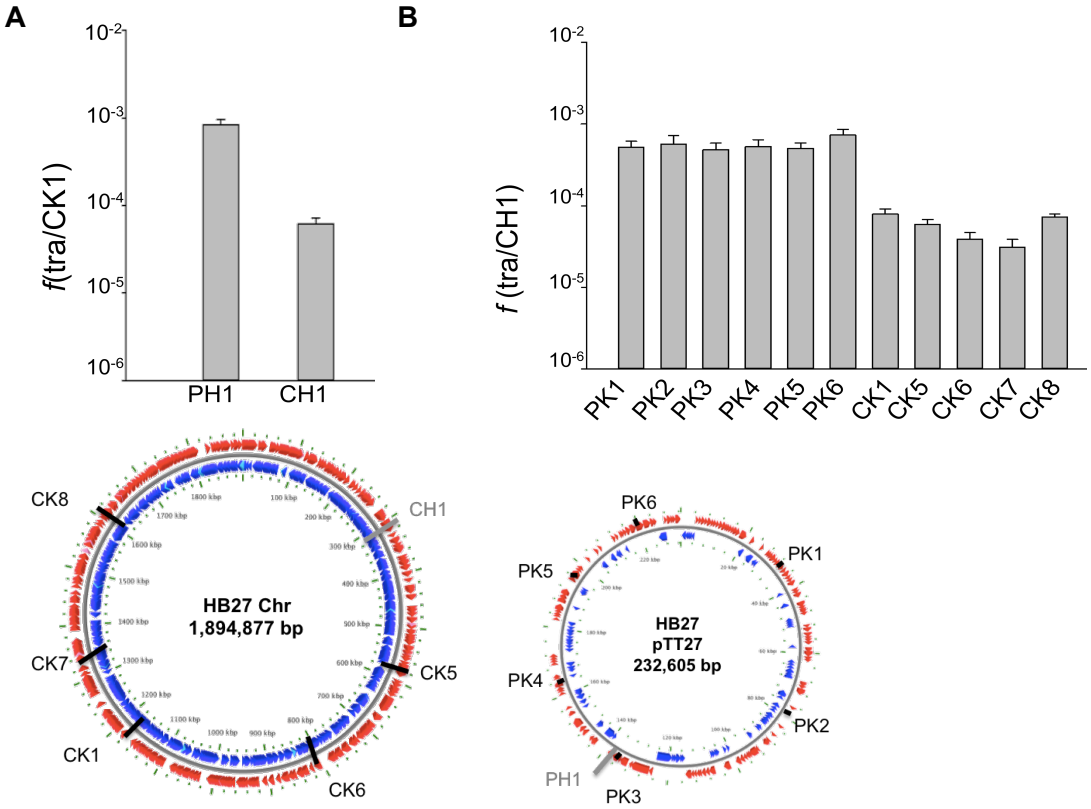


FIGURE 5.3. Chromosomal *vs.* plasmid transfer frequencies among HB27 strains. **A**. Bars represent transfer frequencies of transconjugants per chromosomal CK1 cells obtained from matings between CK1 and Hyg^R chromosome (CH1) or megaplasmid (PH1) labelled strains. Error bars correspond to the mean standard deviation (n=9). **B**. Bars represent transfer frequencies of transconjugants per Hyg^R chromosomal CH1 cells from matings between CH1 and either chromosomal (CK1, CK5-8) or plasmid (PK1-6) Kan^R resistant strains. Error bars correspond to the mean standard deviation (n=9). Maps enclosing the genetic localization of the markers are provided too.

In order to analyse whether there were differences in the transfer of genes associated to their position in the genome, we conveyed a battery of matings involving CH1 with several Kan^R strains, labelled in either the chromosome (CK1, CK5, CK6, CK7, CK8) or in the pTT27 megaplasmid (PK1, PK2, PK3, PK4, PK5, PK6) (Fig. 5.3.B). It was remarkable that transfer frequencies of all pTT27-associated were, at least, one order of magnitude higher than any of those involving chromosomal markers. Therefore, we concluded that there is a strong preference for transfer of genes associated to the megaplasmid compared to chromosomal genes independently of their positions in the respective genetic elements (~9-10 fold; *p*-value< 0.01).

Moreover, we also noticed that there was a rather homogeneity among the transferability of the pTT27 associated genes (*p*-value: 0.634), hindering the inference of an order of transfer in the megaplasmid. Conversely, there was a large variability among the chromosomal-labelled derivatives (*p*-value< 0.01), reporting lowest and highest transfer frequencies for CK7 (*TTC1415::kat*) and CK1 (*TTC1211::kat*), respectively. Implementation of several statistical tools did not provide enough significance to infer a sequential order of transfer for such chromosomal markers.

5.3.3. Conjugation is widespread among *Thermus* strains

In order to analyse the inter-strain lateral inheritance of the denitrification island we mated several Hyg^R and Kan^R derivatives from distinct *Thermus* spp. Matings involving the nitrate respiring HB27d derivative PH20 (HB27d *ΔnirS::hyg*; Table 3.1) with the aerobe CK1 provided Hyg^R+Kan^R nitrate-respiring transconjugants. Besides, strains such as SG0.5 and PRQ25, which were refractory to transformation[48], were capable of transferring the *nar* cluster to the aerobe CK1 strain, providing anaerobically growing Kan^R transconjugants.

In order to fully evaluate the capacity of this inter-strain transfer we crossed HB27 PH1 derivative with different *Thermus* spp strains, chromosomally labelled with the *kat* gene at the highly conserved *pyrE* gene (Table 3.1, Chapter 3). As shown in figure 5.4, broad inter-strain transfer was observed, demonstrating that conjugation was a widespread property amongst *Thermus* strains.

The fact that conjugation was not exclusive of the HB27 strain enabled the distinction of the putative donor and acceptor of the matings, according to the different electrophoretic membrane pattern profile shown by the different *Thermus* strains. Parenthood analysis of some of the grown transconjugants (with both antibiotic resistances) from these mates (HB8, SG0.5, NARI and *T. scotoductus* Kan^R derivatives crossed with HB27 PH1 strain) revealed that, in most cases, transconjugants exhibited

identical membrane pattern profiles to that of HB27, thus, identifying HB27 strain as a universal recipient (data not shown).

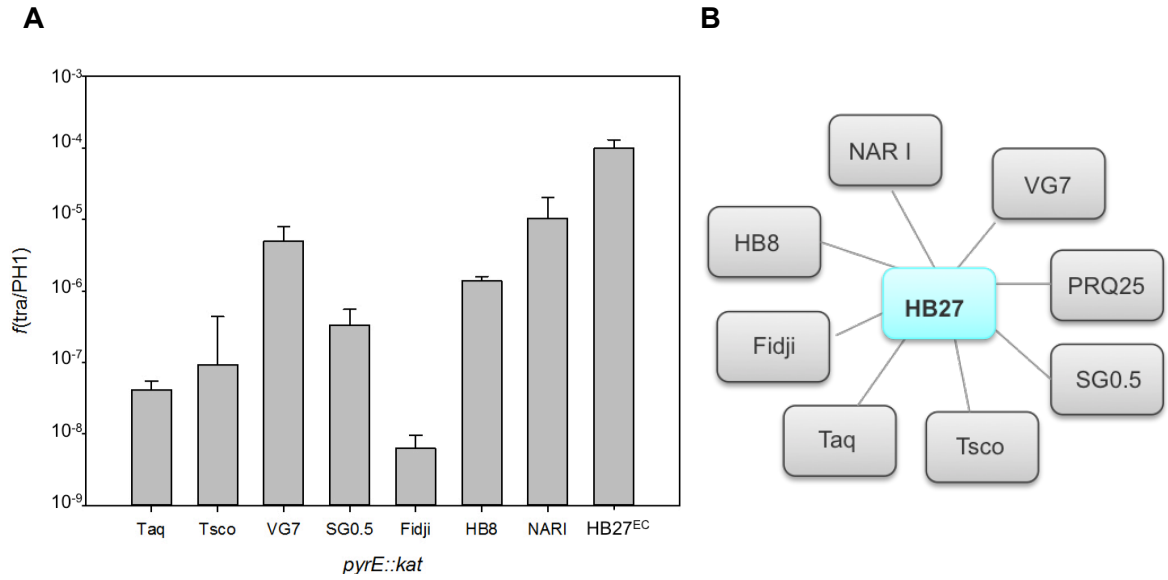


FIGURE 5.4. *Thermus* spp. perform inter and intra-strain conjugation-like transfer. A. Transfer frequencies of PH1 mated to different *Thermus* strains Kan-labelled in the chromosomal loci of *pyrE*. **B.** Sketch representing the different DNA exchanges between *Thermus* spp.

5.3.4. Transformation machinery is required for conjugation

Similarity of components of the competence machinery to proteins involved in conjugation in other bacteria lead us to question whether some of them could display a role in conjugation. Theoretically, if the competence apparatus was to be also the conjugal secretion machinery, a mutant in competence would be impaired in conjugation. Therefore, we designed a battery of experiments in order to investigate the implication of the transformation machinery in conjugation.

To examine a putative relationship between them, we employed a battery of Kan^R mutants of components from the transformation apparatus (CK11-21) in mating experiments with the competent Hyg^R strain CH1. As shown in figure 5.5.A, conjugation was not cancelled in these mutants. However, the large variability among transfer frequencies hinted drawing any further conclusion. Nonetheless, high number of transconjugants were detected in many of the matings, being even more efficient than those measured for the CK1 control (CK12, CK14 and CK15 strains, impaired in PilQ, the prepilins PilA1, 2 and 3 and ComEA, respectively). Besides, transfers were independent of piliation, as certain strains such as CK21, were impaired in competence but still piliated[168].

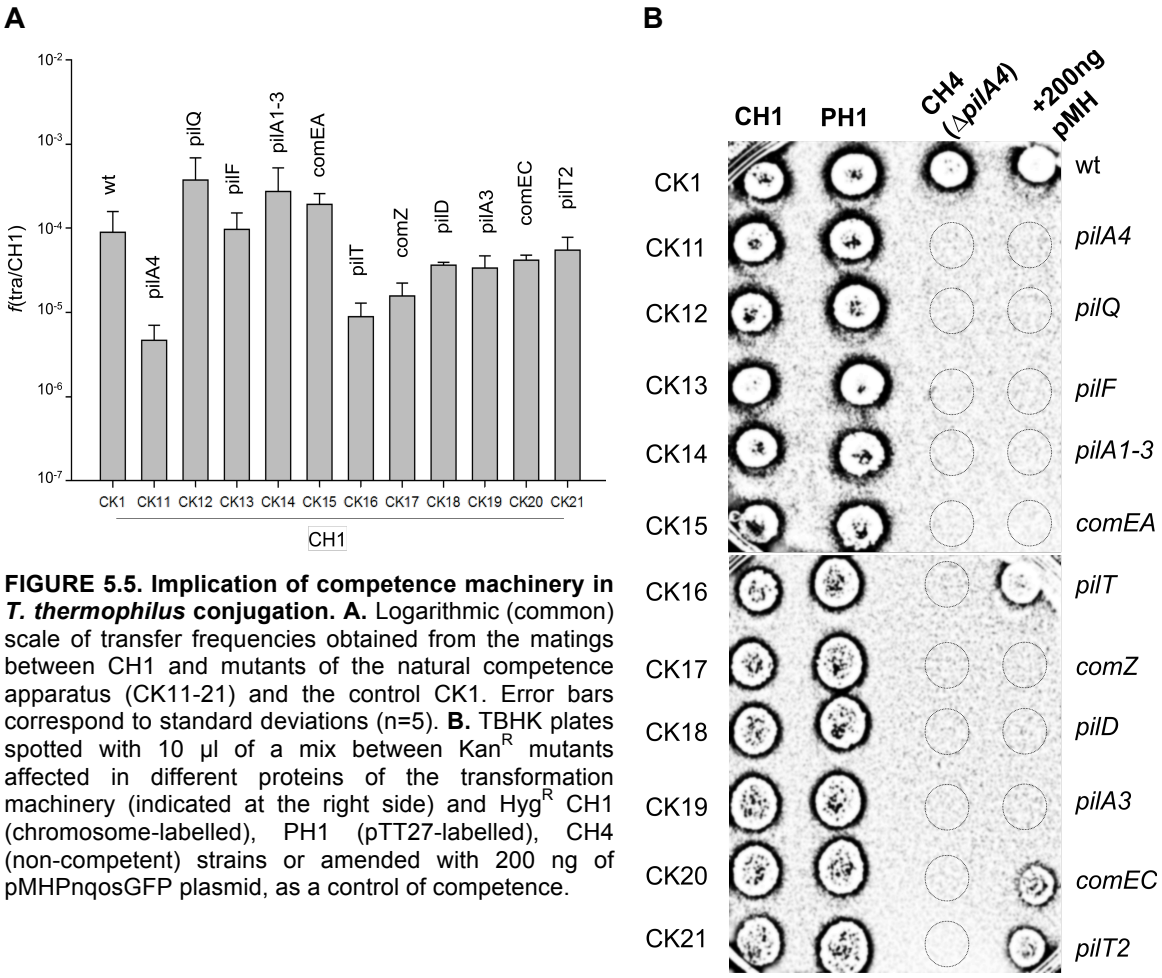


FIGURE 5.5. Implication of competence machinery in *T. thermophilus* conjugation. A. Logarithmic (common) scale of transfer frequencies obtained from the matings between CH1 and mutants of the natural competence apparatus (CK11-21) and the control CK1. Error bars correspond to standard deviations (n=5). **B.** TBHK plates spotted with 10 µl of a mix between Kan^R mutants affected in different proteins of the transformation machinery (indicated at the right side) and Hyg^R CH1 (chromosome-labelled), PH1 (pTT27-labelled), CH4 (non-competent) strains or amended with 200 ng of pMHPnqosGFP plasmid, as a control of competence.

We then monitored pairings involving both mates impaired in competence (Fig. 5.5.B). We observed that when competence-deficient Kan^R strains were mated to another competent-deficient strain (CH4), no transconjugants were obtained in any case. Therefore, the transformation apparatus was somehow involved in conjugation. Besides, some of the mutants were completely unable to uptake DNA whereas others showed detectable transformation (*i.e.*, CK16, Fig. 5.5.B). Altogether, we could conclude that all the components of the competence machinery assayed were required for conjugation in one the mates.

We then examined which of the mates, required a functional competence machinery. As parenthood analysis was not possible between isogenic cells, we constructed a NARI non-competent strain (CH3). The already mentioned phenotypical differentiation between NARI and HB27 at the membrane protein patterns enabled their discrimination, and thus, the identification of parenthood and directionality of the transfer. Hence, *pilQ::hyg* mutant derivatives of the HB27 and NARI strain (CH2 and CH3, respectively) were used in spot mating assays with Kan^R competent HB27 and

NARI strains (CK1 and CK3, respectively) and with competence-deficient strain CK11 (HB27) were mated to CH2 and CH3.

As shown in figure 5.6.A, crosses involving both mates impaired in competence (CK11*CH2; CK11*CH3) did not provide any transconjugants (spots 10-a and 11-a). In contrast, matings CK1*CH3 and CK3*CH2, where the Kan^R strain were competent, produced offspring on TBHK plates (spots 10-b, 11-b, 10-c and 11-c). Transformation controls capacity were also performed, being positive for CK1 (HB27) and CK3 (NARI) (spots 12-b and 12-c) whereas CK11, CH2 and CH3 were unable to uptake plasmid DNA (spots 12-a, 10-d and 11-d). Membrane protein profiles from transconjugants were analysed by electrophoresis and compared to their parental strains (Fig. 5.6.B). In all cases, transconjugant profile fitted that exhibited by the competence-proficient mate: NARI for spots 10-b and 11-b and HB27 profile for spots 10-c and 11-c. These results evidenced that, during mating, cells impaired in competence were able to donate DNA but not to receive it, indicating, thus, that recipient must carry a functional transformation apparatus.

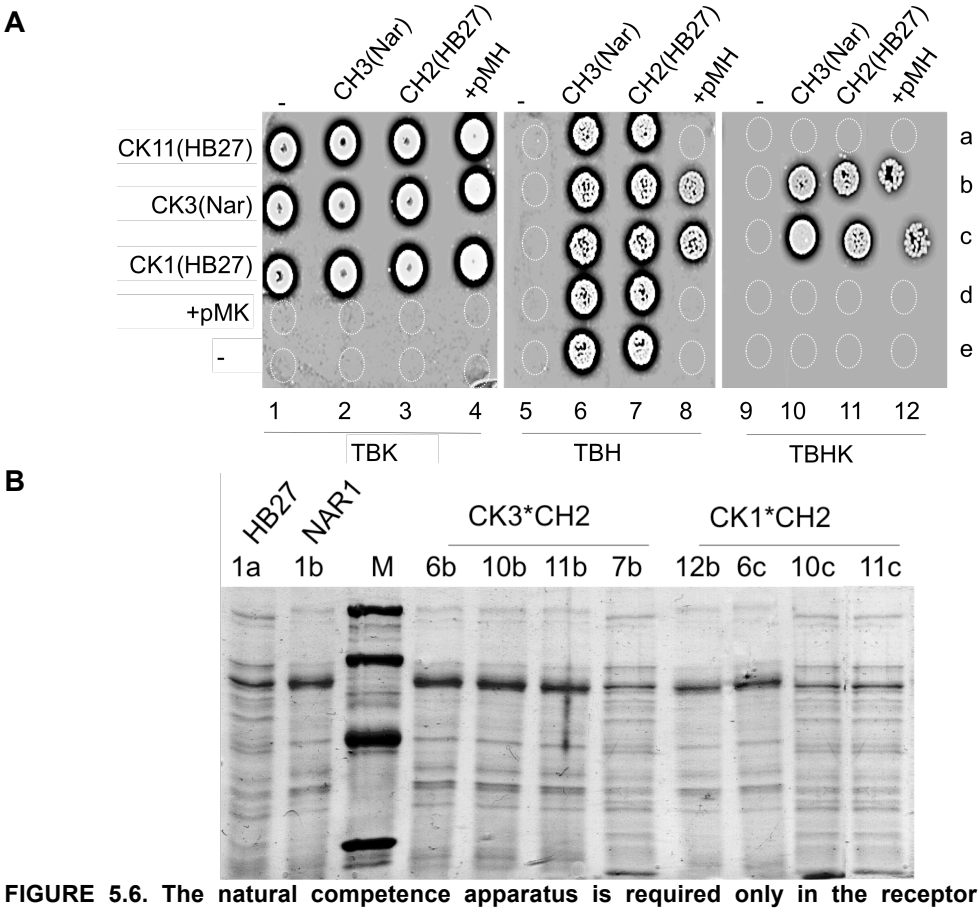


FIGURE 5.6. The natural competence apparatus is required only in the receptor mate. A. Transformation-deficient derivatives of HB27 (CK11, CH2) or NARI (CH3) were mated with competence-proficient derivatives (CK1(HB27 type) and CK3 (NARI type).TB agar plates spotted 10^7 cells of Kan^R strains (vertical lane) topped with either no cells (-), the indicated Hyg^R counterpart derivative (horizontal lane) or 200 ng of purified plasmid DNA (pMK184 and pMHPnqosGFP, respectively). **B.** SDS-PAGE gel showing membrane cell protein extracts from spots of two parental strains (1a, 1b) and transconjugants corresponding to spots of matings shown in panel A.

These results suggested the existence of two active independent mechanisms driving the conjugation-like process: one involved in the pump of DNA (*push*) and another focused on the capture of that DNA (*pull*) (Fig. 5.7). Therefore, the model proposed, so-called "*push-pull*" model, involves a first step where a putative donor actively pushes DNA to the receptor cell in a competence-independent fashion. It is followed by a second step, which requires an active receptor pulling the transferred DNA in order to ease its passage to the cytoplasm, mediated by the competence DNA transporters. Hence, both mates have an active role in the DNA transfer, contrary to classical conjugation model, where the recipient cell remains passive awaiting for DNA transfer[155].

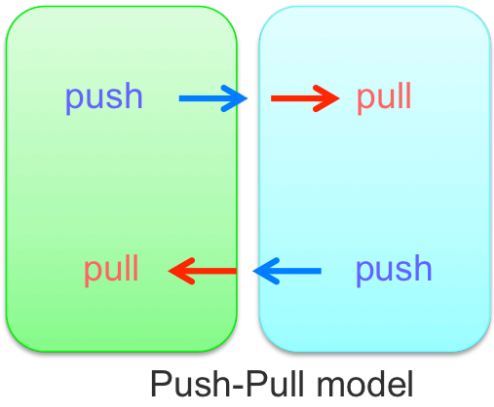


FIGURE 5.7. Two-step *push-pull Thermus* conjugation model. A two step mechanism is proposed to be mediating cell-to-cell lateral transfer of DNA, requiring active participation of both mates. The release of DNA from the donor cell is independent from the competence apparatus, which its presence and functional state is compulsory in the receptor cell.

The nature of this hypothetical *pushing* mechanism is analysed in Chapter 6.

Chapter 6

Exploration of DNA PUSHING during *Thermus thermophilus* conjugation

CHAPTER 6: EXPLORATION OF DNA PUSHING DURING *Thermus thermophilus* CONJUGATION

6.1. SUMMARY

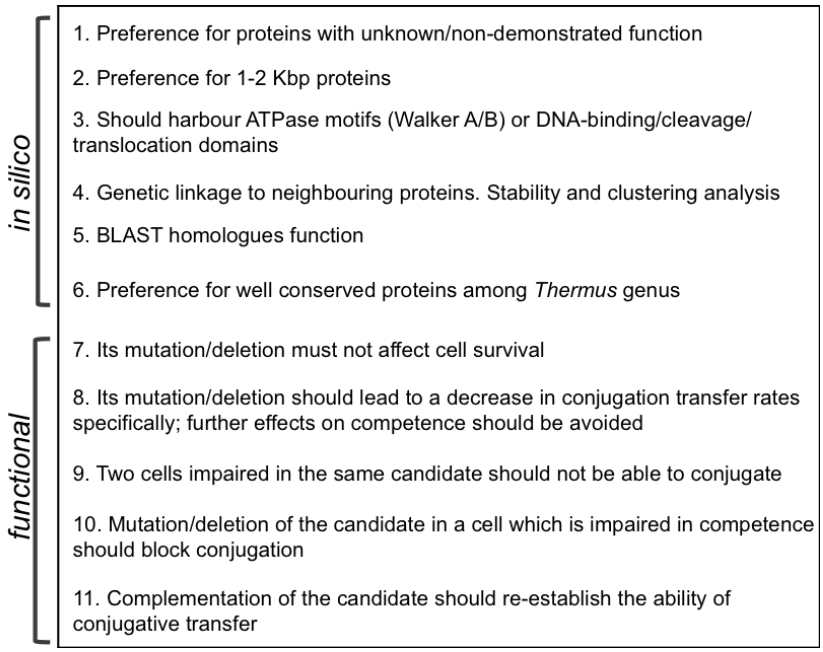
Exploration for proteins involved in DNA donation during conjugation is approached in this chapter. Candidates involved proteins harbouring ATPase or/and DNA-binding motifs. We prioritized proteins with no annotated defined function as well as those where the annotated function lacked experimental demonstration, altogether enfolded as orphan proteins. *In silico* screening provided an output of 23 candidates which were subjected to single recombination mutagenesis in order to check their putative enrolment in conjugative machinery. Just 6 proteins showed actual effects on conjugation, thus, owing a deeper phenotypic analysis. Clean deletions of two of these candidates showed complete blockage of conjugative transfer. An homologue to DNA helicases of the Ftsk-HAS family was found to be specifically involved in DNA donation. Characterization of this protein revealed that it was a functional hexameric ATPase able to bind to the membrane and localised at the cell poles of the cell. Akin, an homologue to helicases of the HerA family homologue conserved among *Thermus* genus, was shown to be required in conjugation and disclosed also effects on natural competence and DNA repair.

6.2. BACKGROUND

In Chapter 5, bacterial conjugation driven in *Thermus* strains has been proposed to follow a two step "*push-pull*" model. Despite no homologues to any of the well conserved proteins involved in standard conjugation were found, we demonstrated that, for conjugation, the competence apparatus was required in the receptor cell but not in the donor. This fact suggested that, if the model was true, there should be an independent mechanism from the competence apparatus accomplishing the pushing of the DNA out of the donor cell. We believed that this task would require a great supply of energy, thus, involving putative ATPases, like it occurs in other unconventional systems such as *Streptomyces*[197]. Besides, any protein which could interact with DNA would have chances to be virtually involved in this initial step of conjugation, mimicking the relaxases or the assisting factors identified in standard models. In fact, more than a fourth of HB27 genome encodes orphan proteins among which we found several of

them harbouring any of these motifs, thus, preliminarily candidates to be involved in conjugation.

In order to search for putative proteins involved in conjugation, a genetic library constructed by insertional random mutagenesis was built in pK18. We gathered 39,721 clones containing around 1Kbp of HB27 genomic DNA which were intended to be transferred to *Thermus* host. However, as commented in Chapter 5, due to the bidirectionality featuring *Thermus* conjugation, any study of candidates involved in DNA donation during conjugation had to be performed on a non-competent background in order to limit the search to the failure in DNA donation. Therefore, a *ΔpilA4* background was required in such library for the search of "*push*" mutants. Unfortunately, the low efficiencies of the electroporation assays (7.28 ± 3.47 $*10^{-9}$) in such genetic background circumvented the execution of this approach. Therefore, we developed an alternative *in silico* screening. Bioinformatic package included BLASTs, multiple sequence alignments and information deposited at the PFAM database, among others. Several statements listed beneath defined the outline of the exploratory analysis of candidates and thus directed the pipeline of the screening, laid on the basis that conjugation should require a prevailing energy supply. Thus, candidates were screened according to the following criteria:



The following sections enclose the results of this screening.

6.3. RESULTS

6.3.1. Screening of potential actors in *T. thermophilus* conjugation

First, 85 pFAMs were initially selected including the ATPases families as well nucleases, helicases and recombinases, for instance. Within the *T. thermophilus* HB27 proteome, 75 putative proteins were enclosed in 42 out of the initial 85 pFAMs. The majority of these candidates were orphan proteins, as expected (Fig. 6.1), but others, such as RecA or RepA, had been demonstrated to display vital functions for the cell and were discarded. Therefore, a discrete study of the positive hits was performed according to the criteria shown above. To the 52 remaining candidates we added 7 proteins more, found by manual inspection across the HB27 genome and considered relevant for the screening according to our previous experimental data (Table S6.1). Principles such as the size or the presence of the aforementioned motifs eliminated remnants of pseudogenes, whereas BLAST hits and pFAM features prompted discarding putative proteins involved in other metabolic pathways. The genomic context of the gene encoding each candidate provided priceless information about its potential clustering within an operon (which was checked by Operon software) as well as whether it was located in a variable region, thus, sensitive to genome rearrangements. This information was relevant in terms of protein functionality and its conservation across *Thermus* genus.

As shown, 23 proteins distributed in 13 pFAMs passed the individual *in silico* analysis and thus, each of them was subjected to insertional mutagenesis in order to perform an individual *in vivo* study to address their potential phenotypical effects on conjugation.

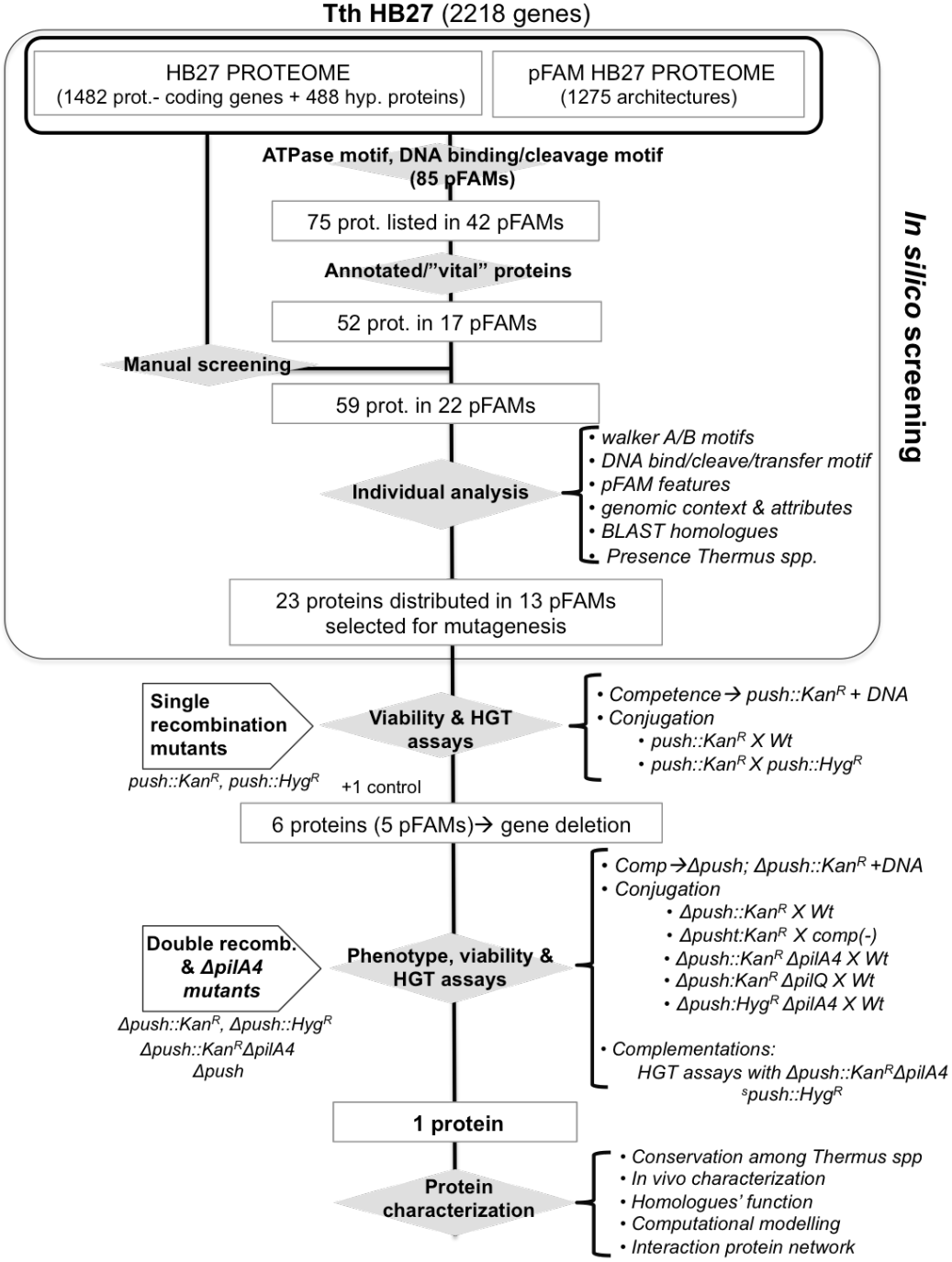


FIGURE 6.1. Flow chart screening potential proteins involved in *T. thermophilus conjugation.* The pipeline includes all proteins studied as potential candidates for DNA donation during *T. thermophilus* conjugation as well as a control, the *gdh* protein (*TTC1211*). Tests developed are listed in the right side.

Single recombination mutants were obtained by transforming the HB27 strain with suicidal vectors pK18 and pH118 harbouring an internal region of each candidate (see Chapter 3 for further details). Each of the Kan^R mutants was subjected to competence tests. Also, conjugation assays were performed by crossing these Kan^R derivatives with its Hyg^R counterparts in a 1:1 ratio, in presence of DNase I. These HGT approach would highlight those candidates were DNA could not be exchanged by

conjugation but were still competent. Viability of the mutants was also addressed. The main results obtained are summarized in table 6.1, which also comprises details of each candidate protein included in the assay.

TABLE 6.1. Characterization of single recombination mutants. Identification of the candidate referred to its encoding locus (Gene bank ID in brackets), characterization (size in *aa*, pFAMs and homologues identified by pBLAST) and results from HGT assays: COMP refers to competence and CONJ to conjugation assays (-:none; +: low; ++: standard; +++: high transfer). Viability (VIAB) was ranked similarly (+: low, ++: standard, +++: high), compared to the control values from CK1 (*gdh::kat*).

LOCUS	AA	PFAM	HOMOLOGUES	COMP	CONJ	VIAB
TTP081 (AAS82411.1)	308	DUF4143 (PF13635.1)	ATPase (*Thermus spp.*; *Meiothermus spp.*)	++	++	++
ttp084 (AAS82414.1)	322	CbiA (PF01656)	Par A (*Thermus spp.; Meiothermus spp.*)	++	++	+
TTP085 (AAS82415.1)	463	RPA (PF10134.4)	RepA (*Thermus spp.; Meiothermus spp.*; *Deinococcus spp.*)	++	++	++
TTP128 (AAS82458.1)	1649	Helicase_C_2 (PF13307)	ATP-binding helicase, RecQ family (*Thermus spp.; Meiothermus spp.*)	++	++	+
TTP140 (AAS82470.1)	446	AAA_14 (PF13173)	ATPase AAA (*Pseudomonas spp.*)	++	++	++
TTP145 (AAS82475.1)	443	RPA (PF10134.4)	No homologue.Putative ATPase, RepA (*Meiothermus spp.)*	+	+	+
TTP191 (AAS82521.1)	551	UvrD_C_2 (PF13538); UvrD/REP helicase (PF00580)	DNA/RNA helicase (*T. oshimani*), DNA helicase (*Cupriavidus spp.*)	++	+++	+++
TTP208 (AAS82536.1)	321	XerD_COG4974; Phague_int_SAM_1	recombinase XerD (*Thermus spp.; Firmicutes bacterium*)	++	+++	+
TTP211 (AAS82539.1)	406	AAA_14 (PF13173)	ATPase AAA+ (*T.aquaticum, T.islandicus, Candidatus spp.*); GTP-binding protein (*D. camini, D. oleovorans*)	++	++	+
TTC0147 (AAS80495.1)	576	AAA_10 (PF12846)	HerA helicase, ATPase (*Thermus spp.*; *Meiothermus spp.*; *Deinococcus spp.*)	++	+	+
TTC0474 (AAS80822.1)	867	AAA_10/22; Ftsk-Y; DUF 87; TRWB _AAD_BIND; T2SE	Cell division protein Ftsk (*Thermus spp.; Meiothermus spp.; Deinoccoccus spp.*)	+	+	+
TTC0638 (AAS80986.1)	857	UvrD_C (PF13361); UvrD-helicase (PF00580)	RecB, ATP-dependent DNA helicase (*Thermus spp.;Meiothermus spp.;M. hydrothermalis*)	++	++	++
TTC0656 (AAS81004.1)	293	Prim-Pol (PF09250)	DNA primase/polymerase (*Thermus spp. ;M. ruber*)	++	++	+
TTC0893 (AAS81237. 1)	728	DEAD (PF00270); Helicase_C (PF00271)	DEAD/DEAH box helicase (*Thermus spp.; Meiothermus spp.;M. Hydrothermalis*)	++	+++	++
TTC1415 (AAS81757.1)	368	T2SE (PF00437)	PilT2, PilU, twiching motility (*Thermus sp; Meiothermus spp; O. profundus*)	+	+++	+++
TTC1429 (AAS81771.1)	381	NurA (PF09376)	NurA nuclease (*C. aerophila; T. scotoductus; Chloroflexus spp.*)	+++	+++	+
TTC1430 (AAS81772.1)	611	AAA_10 (PF12846); GAT_1	ATPase (*P. methylaliphatogenes; C. aerophila; T. scotoductus; Chloroflexus spp.*) ComB4/HerA helicase (*M. infernorum*)	+	+	++
TTC1621 (AAS81963.1)	321	T2SE (PF00437)	twiching motility protein PilT (*Thermus spp.; Meiothermus spp.; Deinococcus spp.*)	++	++	++
TTC1622 (AAS81964.1)	889	T2SE (PF00437)	PilF, ATP-binding protein (*Thermus spp.; Meiothermus spp.; Deinococcus spp.*)	-	+	++
TTC1836 (AAS82178.1)	116	N_methyl (PF07963)	Cleavage, N-terminus pili involved in secretion GspG (*T. yunnanensis*); comGC (*B. subtilis*)	+	+	+
TTC1839 (AAS82181.1)	150	N_methyl (PF07963); T2SE_G; T2SE_J; PulG	GspG (*T. yunnanensis*); comGA (*B. subtilis*)	+	+	+
TTC1844 (AAS82186.1)	548	T2SE (PF00437)	PulE, PilB. ATPase (*Thermus spp.; M. hydrothermalis; Meiothermus spp.*)	-	+	+
TTC1879 (AAS82221.1)	568	AAA_10 (PF12846)	Ftsk/HerA, ATPase (*Thermus spp.; Meiothermus spp.; O. profundus*)	++	-	++

From these assays we observed several candidates showing promising results regarding their potential role in conjugation. Among them, we selected those showing standard yields of viability. Besides, we initially discarded competence homologues already demonstrated (*i.e.*, *TTC1621, TTC1622* encoding PilT and PilF, respectively) as their function had already been addressed in detail[167], shown to be involved in the pulling step, so we prioritized on candidates for systems independent from the competence apparatus. Constructs of 14 single recombinant candidates were electroporated in *ΔpilA4* cells to avoid them acting as recipients in matings. Each of these non-competent derivatives were mated to the Hyg^R wild type strain CH1, which could only act as recipient under these experimental conditions (Fig. 6.2). Also, each candidate was crossed with CH4, which is impaired in competence (*ΔpilA4*), as a negative control.

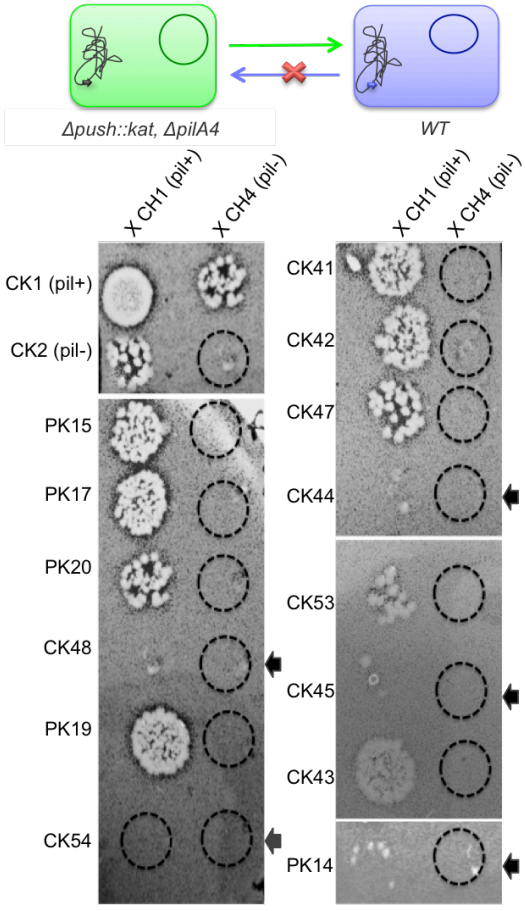


FIGURE 6.2. Experimental search for competence-deficient candidates. Conjugation spot assays involving 10^7 cells of *push::kat ΔpilA4* candidates with either CH1 or CH4 (*ΔpilA4)*. Final candidates, where conjugation is blocked, are highlighted with an arrow. Loci corresponding to each strain can be checked in table 3.1: CK1 and CK2: *GDH*; PK15: *TTP140*; PK17: *TTP191*; PK20: *TTP081*; CK48: *TTC1430*; PK19: *TTP208*; CK54: *TTC1879*; CK41: *TTC0638*; CK42: *TTC893*; CK47: *TTC1429*; CK44: *TTC1844*; CK53: *TTC1839*, CK45: *TTC0147*; CK43: *TTC1415* AND PK14: *TTP145.*

As it can be observed in figure 6.2, 5 derivatives (CK44, CK45, CK48, CK54 and PK14) were defective in conjugation. In addition, a sixth candidate (CK47; *ΔTTC1429::kat ΔpilA4*) was included in further studies, according to its possible co-transcription with another candidate (*TTC1430*). Moreover, previous results from a double mutant *TTC1429-TTC1430* showed significant decreases in conjugative transfer, suggesting, thus, potential involvement in DNA mobilization in association to

another candidate (CK48; *ΔTTC1430::kat ΔpilA4*) (data not shown). Therefore, stable deletion mutants of these candidates through double recombination were obtained and their phenotype examined.

6.3.2. Analysis of candidates involved in DNA donation

Recombinant deletion mutants (*Δpush::kat*) and markerless derivatives (*Δpush)* were obtained (Chapter 3) and a battery of phenotypical assays were performed with these derivatives which results are summarized in table 6.2. Complementation assays with the respective wild type gene were performed too (Fig. S6.1; Annex IV).

As it can be observed in table 6.2, *nurA1* and *repA2* mutants had not a significant effect on conjugation rates. On the other hand, *cptB* and *pulE* mutants showed a ~100 fold decrease in conjugation efficiency. Finally, conjugation was completely abolished in *herA* and *cptA* mutants. These two mutants were complemented and conjugation was recovered (Fig. S6.1; Annex IV).

Apart from the effect on conjugation, we observed that impairment of some of these genes caused multiple phenotypic effects. We were particularly concerned about the decreases in viability in *nurA* and *repA2*, and, in a lesser extent, in *herA, cptB* and *pulE*, which could mask the real role in conjugation. In several cases competence was affected. Greatest reduction was starred by *pulE*, which, according to its genomic context (Fig. 6.3), was enclosed within a putative T2SS harbouring several prepilin homologues. Transformation rates were recovered once the gene was expressed from a plasmid. Other candidates such as *herA* and *repA2*, were affected in transformation too, showing ~15 fold decrease in competence rates. Moreover, except for *cptA*, all candidates showcased a decrease in the yields of viable cells, specially *repA2* and *pulE*, which evidenced a ~100 fold decrease. Reduction of viability was also encompassed with lower survival after a “heat shock” or higher UV-sensitivity shown by *pulE, repA*2, *herA* and *cptB*. Besides effects on other physiological parameters such as twiching-motility and adherence, or biofilm formation were observed among the mutants in different degrees. In contrast, deviations from control values were not as sharp in adherence and twiching motility.

All in all, from the analysis of the data enclosed in table 6.2, we could identify two of the candidates, *cptA* and *herA* as the most relevant aspirants to be involved in conjugation. Therefore, a deeper analysis of these two candidates was carried out.

TABLE 6.2. Characterization of the candidate proteins involved DNA donation during *T. thermophilus* conjugation. *gdh* strain is presented as a control and *nurA1* was included according to its potential relation to *cptB.* Numerical results were a mean of at least 5 independent experiments (mean±SD) and n.d referred to no data available. Transformability data was referred to 10-20 ng DNA. UV sensitivity was estimated compared to the control HB27wt: - (no difference), ++ (completely different). Survival after T° raise referred to the CFU per ml grown after an overnight incubation at 79.2 ºC. Twiching-motility and adherence were referred as the diameter (in mm) of the twiching or adherence area, respectively. Subcellular prediction was implemented by the consensus data from Psortb and CELLO databases: IM (inner membrane), C (cytoplasmic), P (periplasmic), OM (outer membrane). Conservation in *Thermus* spp was rated according to sequence alignments. pBLAST threshold was set at an E-value of 10^{-20}: - (just in HB27), + (2-4 strains), ++ (4-6 strains), +++ (> 6 strains).

Genotype	Proposed name	Strain code	Conjugation frequency rates *ΔpilA4* [*f*(tra/HygR)]	Transformation frequency rate [*f*(tra/viable cells)]	Viability [*cfu* Kan^R/ml]	UV sensitivity	Survival after T° raise	Twiching area [mm]	Biofilm formation[Abs_{57_0}/OD_{600}]	Adherence [mm]	Putative subcell. loc	Conservation in *Thermus spp*
ΔTTC1211::kat	Gdh	CK1	2,1E-5 ± 3,8E-6	2,2 E-4 ± 6,0E-5	2,5E9 ± 2,1E8	-	1,8E8 ± 4,3E9	1,2 ± 0,2	3,8 ± 1,0	1,6 ± 0,5	C	+++
ΔTTP145::kat	RepA2	PK26	1,31E-4 ± 2,55E-5	7,3E-6 ± 9,3E-7	3,8E7 ± 1,1E7	+	n.d.	1,1 ± 0.5	5,0 ± 0,9	1,4 ± 0,3	IM	-
ΔTTC1429::kat	NurA1	CK74	5,2E-5 ±3,7E-6	2,8E-5 ±3,01E-6	6,7E7 ± 1,1E7	-	n.d.	n.d.	3,3 ± 0,5	n.d.	C	+
ΔTTC1430::kat	CptB	CK71	1,6E-7 ± 6,8E-7	1,4E-5 ± 9,1E-6	1,4E8 ±4,9E7	+	7,2E5 ± 2,3E5	1,0 ± 0,2	3,0 ± 0,5	1,5 ±0,7	C	+
ΔTTC1844::kat	PulE	CK68	1,3E-7 ± 3,3E-8	8,1E-7 ± 4,0 E-8	6,1E8 ± 2,1E8	+	2,0E6 ± 7,9E5	0,8 ± 0,1	4,7 ± 1,1	1,2 ± 0,3	C	++
ΔTTC0147::kat	HerA	CK65	0	8,1E-6 ± 1,8E-6	7,5E7 ± 1,9E7	+	9,1E6 ±7,2E5	1,0 ± 0,1	5,9 ± 0,8	1 ± 0,2	C	+++
ΔTTC1879::kat	CptA	CK55	0	4,4E-5 ± 1,0E-5	2,2E9 ± 7,6E8	-	1,0E8 ± 5,6E7	1,1 ± 0,2	5,2 ± 1,0	1 ± 0,3	C	+

6.3.3. Bioinformatic analysis of the candidates involved in DNA donation during *T. thermophilus* conjugation

In parallel to the phenotypical study, bioinformatic analysis of the genomic context of each of the 6 candidates was carried out (Fig. 6.3).

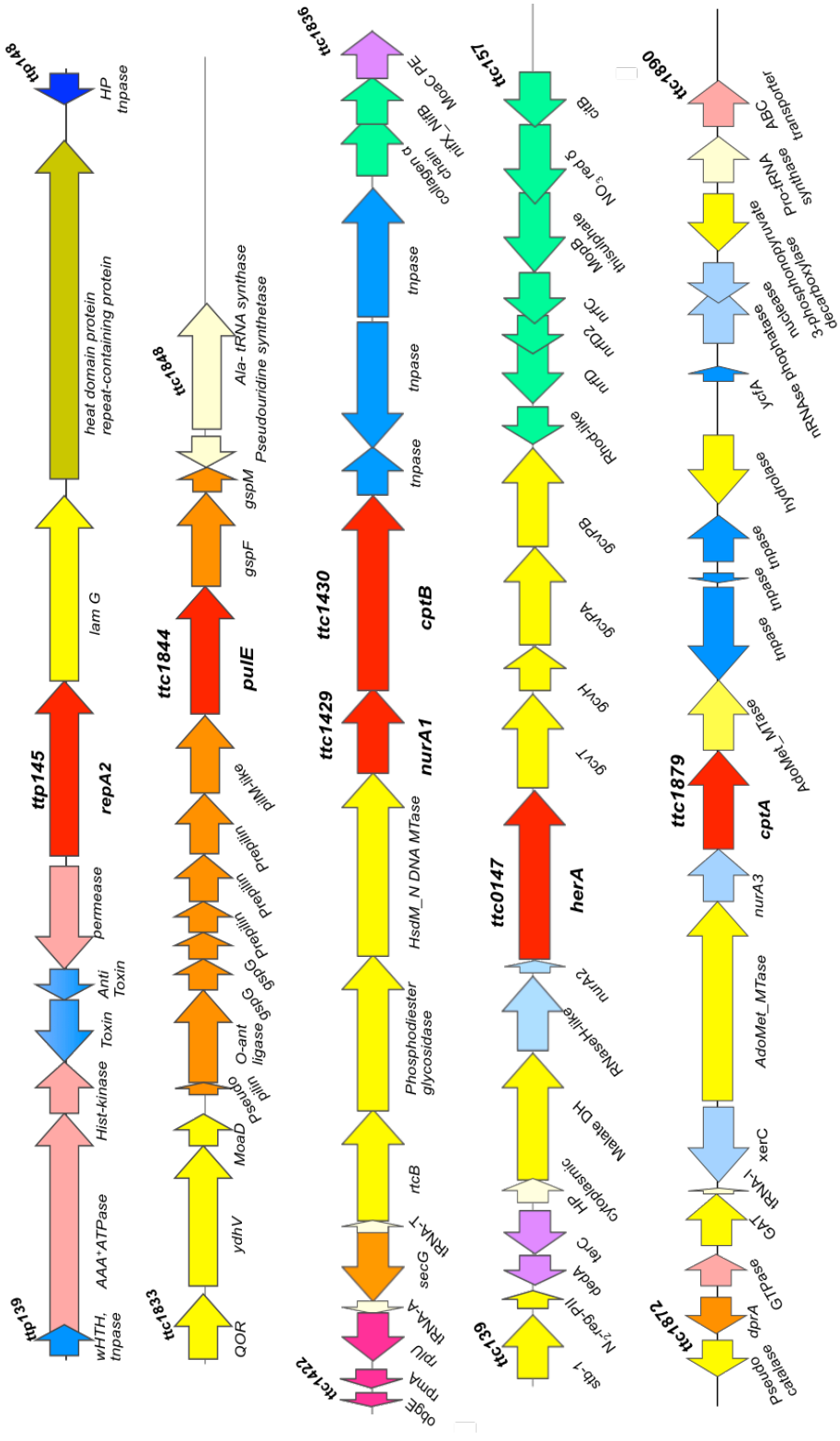


FIGURE 6.3. Genomic context of the 6 candidates for DNA donation during *T. thermophilus* conjugation. Each of the candidates is shown in red, proteins labelled or presenting DNA- modifying motifs are represented in light blue. Transposases in dark blue; putative proteins addressed to metabolic pathways in yellow and putative transporters in pink. Putative respiratory proteins are represented in green, integral predicted membrane proteins in purple and t-RNAs in light yellow

As observed in figure 6.3, several candidates were encoded within fairly conserved regions and shared a similar genomic context. In fact, *cptA, cptB* and *herA,* included a putative *nurA*-like nuclease immediately upstream the candidate. Interestingly, these three candidates were clustered in the same PFAM: AAA_10 ATPases suggesting similar enzyme activities(Table S6.1). Other candidates, such as *cptA, repA2* and *cptB* harboured transposase-like motifs downstream, suggesting possible mobilization. Incidence of these latter candidates across *Thermus* spp was restricted to some of the sequenced strains or just detected only in HB27 (Table 6.2). Just *repA2* was localized in the pTT27 megaplasmid, which harbours the lowest synteny in the genus. In addition, this was the only candidate predicted to be located in the inner membrane, whereas the remaining candidates were predicted to be soluble cytoplasmic proteins (Table 6.2).

6.3.4. Genomic context of *cptA* and *herA*

Apart from this shared phenotype in conjugation, *cptA* and *herA* presented several similarities, such as the predicted size (64 KDa), the presence of protein motifs (AAA_10 domains) and high homology between these two proteins (E-value 9E-56, 31 % identity). HerA is conserved in HB8 strain (*TTHA0522*, 99 % identity), where it has been described as an archaeal-like DNA helicase involved in DNA repair[58] whereas CptA is not present in this strain.

Genetic context of both HerA from HB27 and its counterpart *(TTHA0522)* from HB8 were conserved in all the sequenced different strains of *T. thermophilus* and even encoded in *T. scotoductus* (Fig. 6.4). Surrounding genes include a glycine processing system downstream and metabolic and nucleotide synthesis related genes upstream. Two conserved hypothetic integral membrane proteins were encoded upstream too.

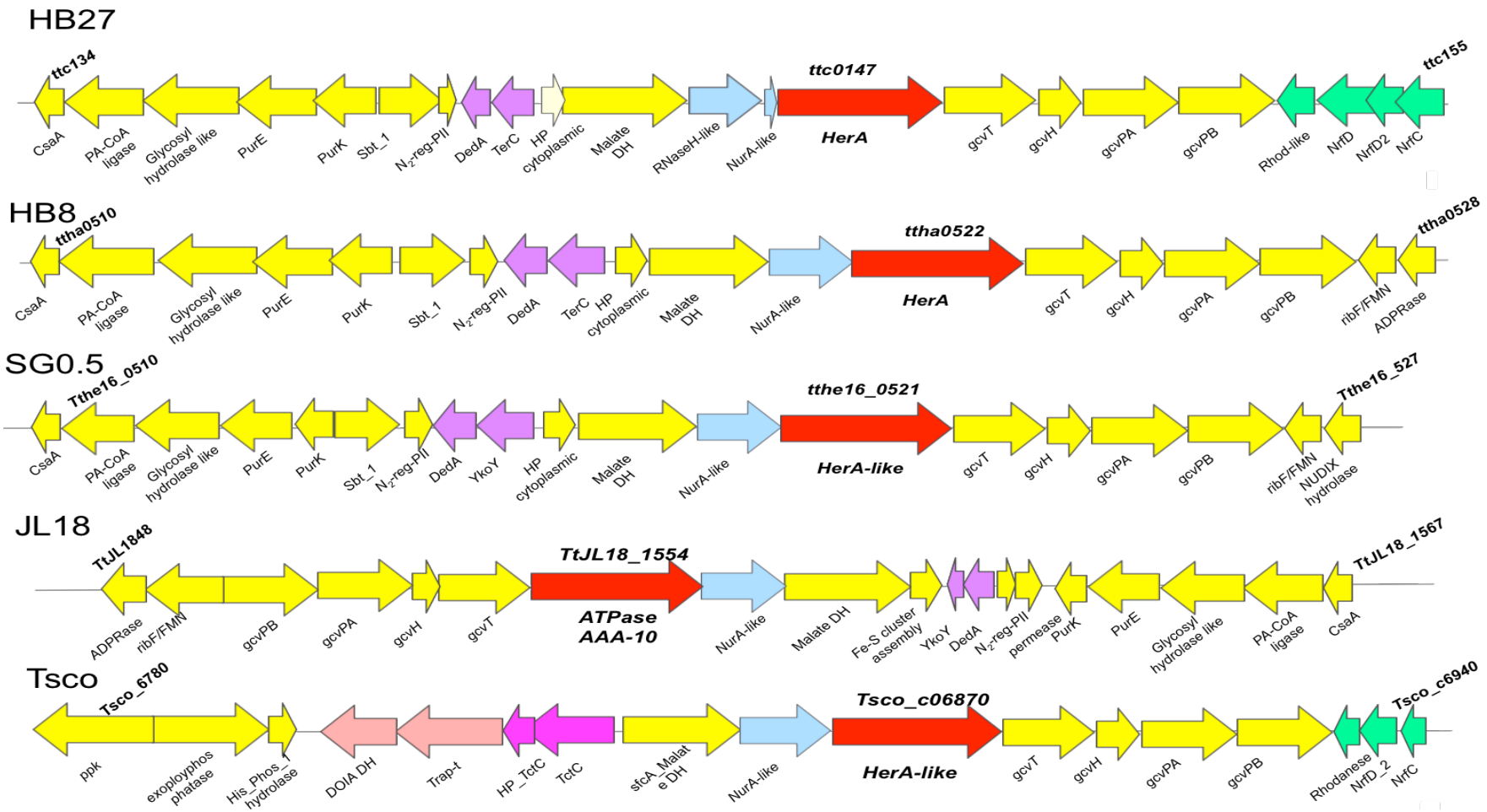


FIGURE 6.4. HerA homologues' genomic contexts. Synteny of the cluster in which the common homologue *TTHA0522* in HB8, was compared to those found in *T. thermophilus* strains HB27 (*TTC0147),* SG0.5JP16-17 (*TTHE16_0521*), JL18 (TtJL18_1554) and *Thermus scotoductus (Tsco_c06870)* cluster. Colour-code legend based on putative functions annotated: each homologue is shown in red, proteins labelled or presenting DNA- modifying motifs are represented in light blue. Transposases in blue; proteins addressed to metabolic pathways in yellow and transporters in pink. Respiratory functions are represented in green, integral membrane proteins in purple and t-RNAs in light yellow

Conversely, TTC1879 locus, coding for *cptA*, presented a different scenario. Despite its singleness in HB27, homologues were found in genomes from other thermophiles (Fig. S6.2; Annex IV). In HB27, upstream the *nurA3-cptA* genes, the gene *TTC1877* encodes a putative S-adenosylmethionine-dependent restriction-modification type II enzyme, highly similar (99 % of identity) to *Tth111*II which cleaves the top strand at N11/N9 downstream the CAARCA (R=A or G) sequence. Interestingly, *CptA* is encoded within a 13-Kbp region of lower G+C content (61 %) than the genome (69 %), edged by IS sequences, a putative phage integrase and a t-RNA. This region also encodes XerC, a host recombinase related to mobilomes and found in conjugative plasmids[213]. Synteny comparison between the HB27 and HB8 genomes supported the acquisition of this region by HGT (Fig. 6.5).

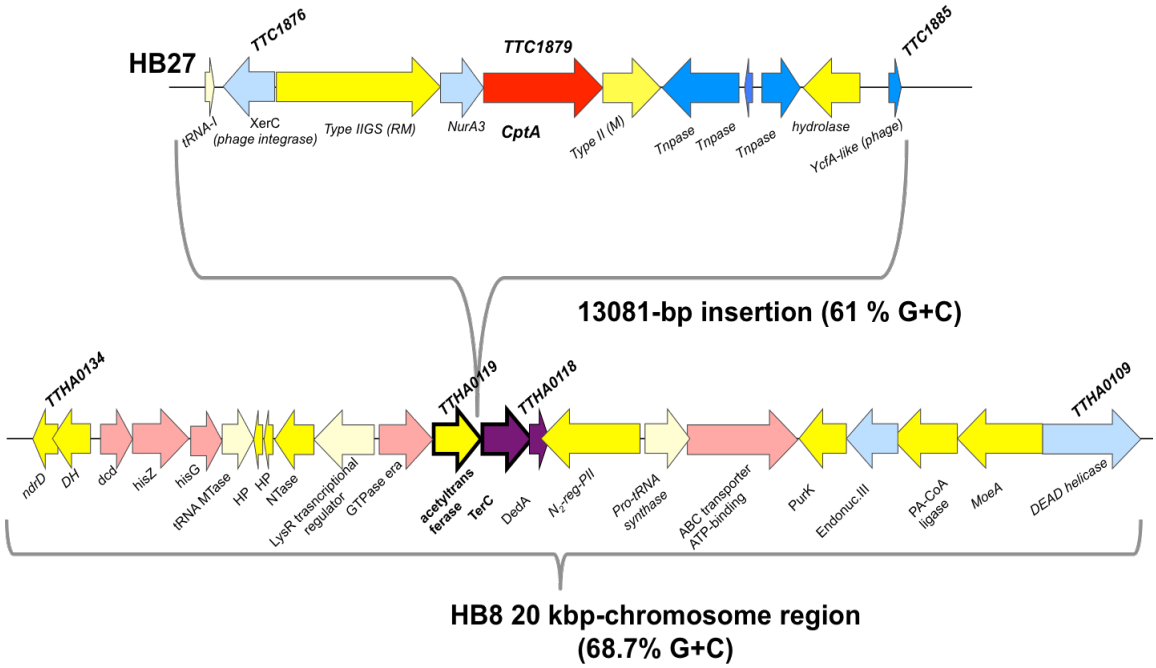


FIGURE 6.5. CptA should had been inserted recently in HB27 genome by HGT. Descriptive representation of the ~13-Kbp region recently inserted in HB8 genome. Colour-code legend based on putative functions annotated is the same as in figure 6.4.

6.3.5. CptA and HerA are required for DNA donation during conjugation

Genetic assays displayed on *cptA* and *herA* mutants evidenced a complete blockage of DNA transfer. In order to ascertain their specific role in conjugation, we compared membrane protein patterns from transconjugants obtained in crosses involving derivatives of strains NARI and HB27, both of which containing homologues of CptA and HerA, as verified by PCR and sequence alignment.

Results showed that transconjugants exhibited the protein pattern corresponding to *cptA* mutant partner (Fig. 6.6). For instance, transconjugants from

matings between CK59 (NARI, CptA$^-$) (Fig. 6.6, lane 1) and HB27 mate (CH1; CptA$^+$) (lane 4) or a non-competent HB27 strain (CH4; CptA$^+$, PilA4$^-$) showed a protein pattern corresponding to the NARI profile (lanes 4 and 5). Conversely, matings between a HB27 *cptA* mutant and the NARI wild type CH3 strain produced transconjugants endorsing an HB27 profile (lane 6). These data confirmed that mutants in CptA could only work as recipients in mating experiments. Equivalent experiments were performed with *herA* mutants with similar results (lanes 8 and 9). Therefore, these data suggested that both proteins were required for DNA donation during conjugation.

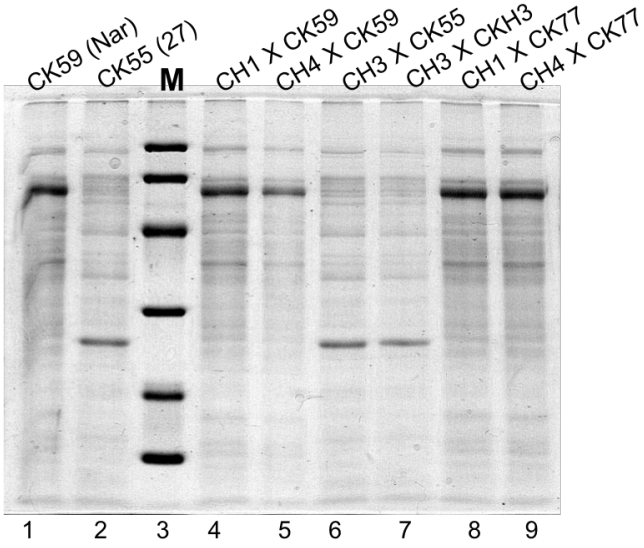


FIGURE 6.6. Unraveliing parenthood in matings. SDS-PAGE of whole membrane proteins from transconjugants grown from matings involving wild type strains HB27 (CH1) or NARI (CH3), as well as a non-competent HB27 strain (CH4) crossed with NARI and HB27 *cptA* and *herA* mutants: *cptA* in NARI (CK59) and HB27 (CK55), and *herA* mutant CK77 (NARI*)*. Strain HB27 CKH3 is impaired in both HerA and CptA. Note that the protein pattern always corresponds to the strain lacking CptA or/and HerA. Lane M, molecular markers 97.4, 66.2, 45, 31, 21.5 and 14.4 KDa.

6.3.6. Localization of CptA and HerA proteins

Localization of CptA and HerA was approached by different methods. *In silico* prediction suggested that both were cytoplasmic (Table 6.2). To check this, Western-blot were performed using antisera against the HB8 HerA homologue, which shared more than 99 % and 31 % of sequence identity with HB27 HerA and CptA, respectively, thus cross-reacting with both proteins.

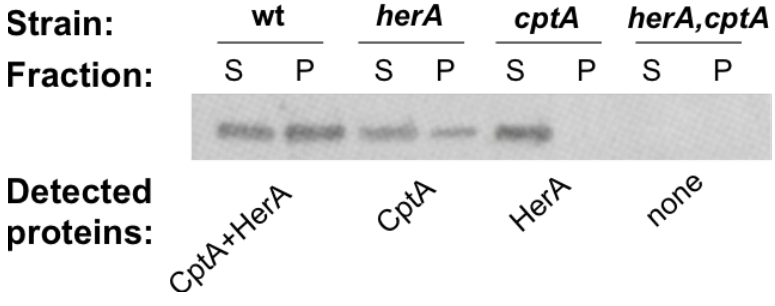
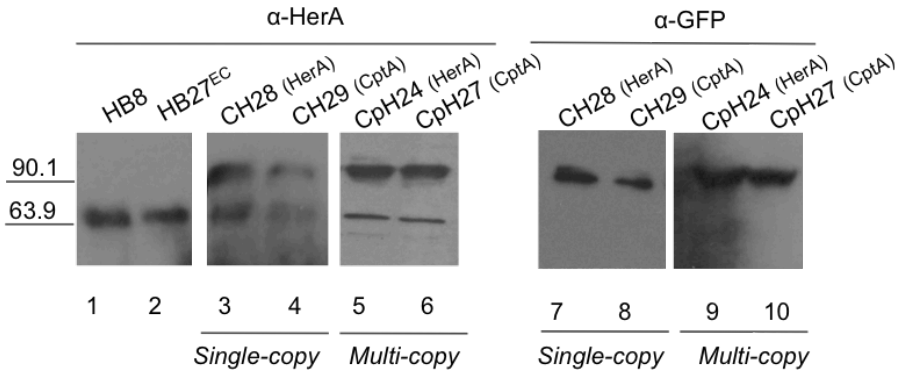


FIGURE 6.7. Subcellular localization of HerA and CptA. Western-blot with α-HerA antiserum of soluble (S) and non-soluble (P) fractions from *T. thermophilus* HB27: Wild type (Wt), CK65 (*ΔherA*), CK55 (*ΔcptA*) and CKH3 (*ΔcptA, herA*$^-$). The proteins detected in each case are indicated underneath.

In the wild type strains, a single band was observed of an apparent size of 64 KDa, in agreement with the predicted size for both proteins (Fig. 6.7). In contrast, double mutant strains impaired in both HerA and CptA (CKH3) did not show any detectable band. In *cptA* (CK55) and *herA* (CK65) individual mutants, a single band was detected (Fig. 6.7), corresponding to HerA and CptA, respectively. However, whereas HerA was only found in the soluble fraction, part of CptA was associated to the particulate one, likely bound to the membrane.

Protein localization was also approached by using C-terminal gene fusions to the thermostable yellow fluorescent protein sYFP[44]. Fusions were obtained either by single recombination with a suicide plasmid (pAB219 and pAB213, for *herA* and *cptA*, respectively) or expressed from a multi-copy vector (pAB175 and pAB224, respectively). Fusion expression of strains harbouring these single and multi-copy constructs (CH28 and CpH24 for *herA* and CH29 and CpH27 for *cptA,* respectively; Table 3.1 in Chapter 3) were tested in Western-blot (Fig. 6.8). Immunoblotting detection by α-HerA and α-GFP antisera was positive in all cases, as shown in figure 6.8. When α-HerA antiserum was employed, samples from HB27 derivatives harbouring single recombinant fusions showed two bands the bigger one (~90.1 KDa) corresponding to the expected size for the fusion of HerA or CptA to the sYFP reporter and a smaller one (~64 KDa) of the size expected for CptA or HerA, respectively (Fig. 6.8.A, lanes 3 and 4). Likewise, α-GFP detection on the same samples proved the fusions by labelling a single band of the size expected for the fusion (lanes 7 and 8). Multi-copy derivatives showed the identical pattern but intensities of the upper bands were higher, according to its multi-copy expression (Fig. 6.8.A, lanes 5-6 and 9 and 10).

A



B

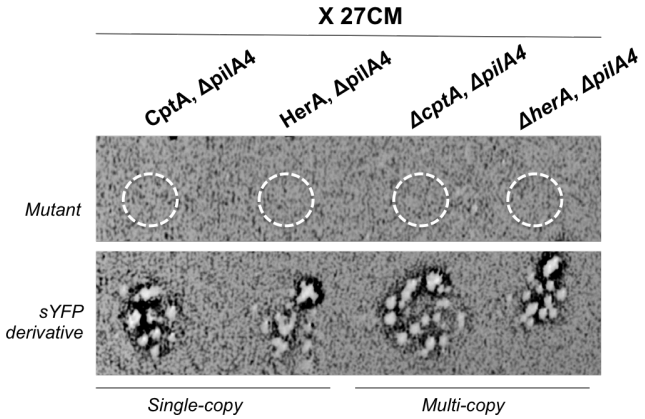


FIGURE 6.8. Functionality assays of fluorescent fusion constructs. A. Immunoblotting detection of GFP reporter fused to the C-terminal region of HerA (CH28 and CH24) and CptA (CH29 and CH27) expressed in single copy and multi-copy vectors, respectively. Native detection of both proteins was positive using α*-HerA* on whole extract proteins from HB8 and HB27^{EC} strains (lanes 1 and 2). Antibodies α-HerA and α- *GFP* were diluted 1:5000. **B.** Conjugation spot assays between HB27 Cm^{R} strain (27CM) and HB27Δ*pilA4* derivatives harbouring CptA-sYFP or HerA-sYFP fusions inserted as single copy in the chromosome or HB27Δ*pilA4* Δ*cptA*/Δ*herA* harbouring multicopy plasmids expressing CptA-sYFP or HerA-sYFP.

All these fluorescent strains showed similar growth yields as the HB27wt strain and were functional in conjugation assays in combination with *pilA* mutations. In this line, electroporation of CK45 (*ΔherA, ΔpilA4*) and CK56 (*ΔcptA, ΔpilA4*) with ectopic fluorescent expression plasmids of HerA and CptA (pAB175 and pAB224, respectively) enabled recovery of conjugation. On the other hand, insertional *ΔpilA4* mutants harbouring single recombinant fusions of sYFP to *herA* and *cptA* were also able to transfer DNA by conjugation (Fig. 6.8.B), thus, showing that the fusions were still functional *in vivo*.

Then, sYFP fusion derivatives were examined by confocal microscopy. As shown in figure 6.9, both proteins exhibited a mainly polar localization whereas expression of the fluorescent protein alone (PpH27 derivative) produced an homogenous labelling of the whole cell (Fig. S6.3; Annex IV). Multi-copy variants confirmed the polar localization, being bipolar in the case of HerA, whereas CptA derivatives showed rather differential expression among the poles, being more intense in one of them. Under the same conditions, HerA showed higher fluorescence intensity than CptA (Fig. 6.9.B), suggesting higher expression levels in agreement with the Western-blot assays (Fig. 6.7).

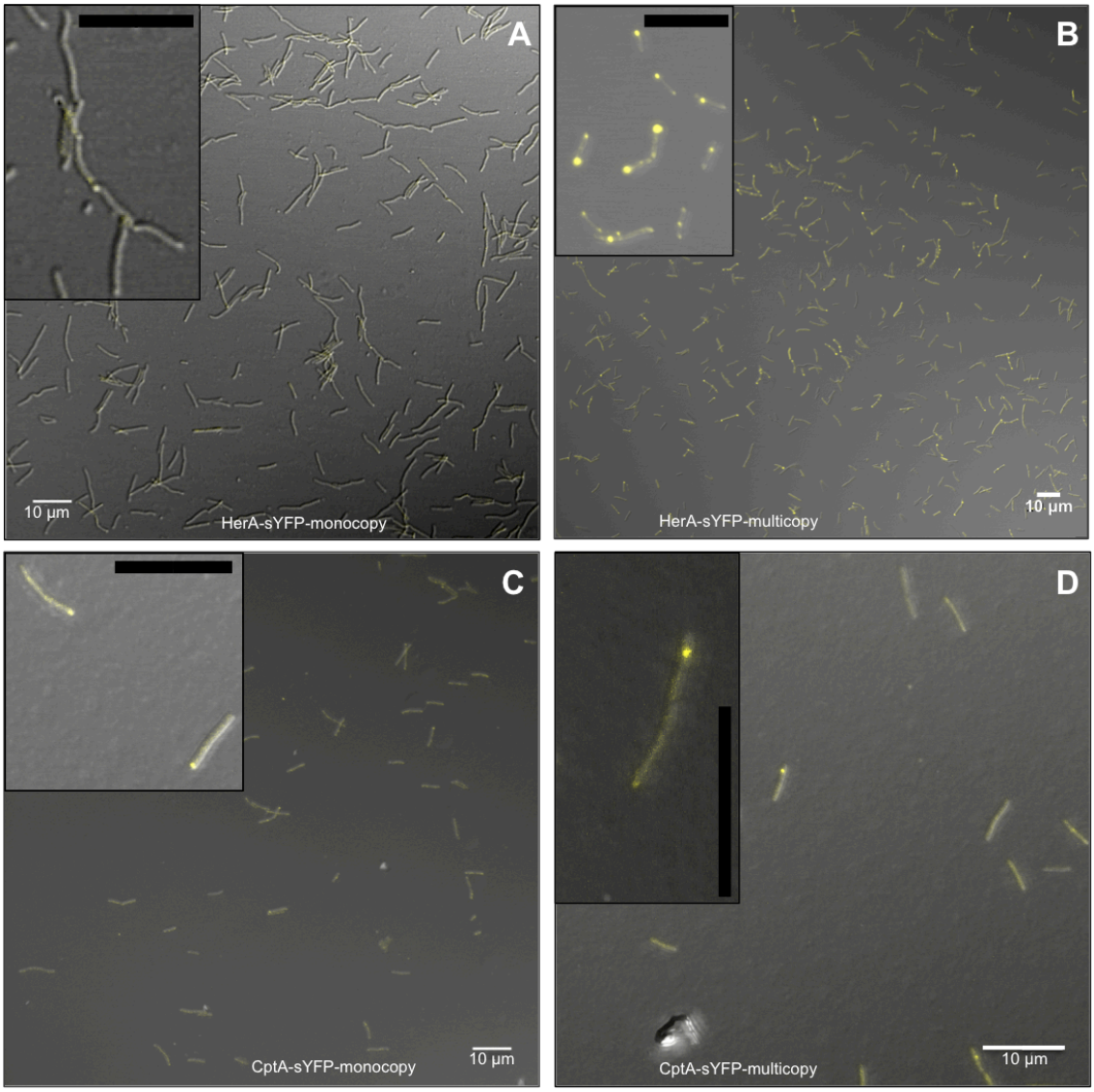


FIGURE 6.9. CptA and HerA are located in the cell poles. Confocal microscopy overlapping images of bright field and yellow channel from HB27 preparations of HerA::sYFP (upper panels, A and B figures) and CptA::sYFP (lower panels, C and D figures). A and C images corresponded to single-copy recombinants, whereas B and D correspond to multi-copy expression from plasmids. Both a general view and a zoomed image were provided in each case.

6.3.7. Synthesis of HerA and CptA proteins

To check if CptA or HerA where expressed differentially among the growth cycle we used the HB27 strains harbouring the single copy fusions to the sYFP reporter (CH29 and CH28, respectively) as specific tags for immmunodetection. Apparently, both proteins were expressed in a constitutive fashion along all the bacterial growth phases (Fig. S6.4 shows CptA case), although levels detected were very low at low OD.

Examination of potential linkage between CptA expression and quorum-sensing signals involved incubations of growing CH27 cultures with filtered supernatants from

stationary phase cultures of the same strain, but no difference in protein expression could be inferred (data not shown).

Inducibility of these proteins under cell contact or in presence of DNA was examined too. Cultures of CH29 and CH28 were grown to final exponential phase (OD_{550} of 0.6 - 0.8) and supplemented with genomic DNA, or with equal cell amounts of an HB27wt strain. Mixtures were incubated and then processed for Western-blot analysis. No significant change on protein expression levels could be detected (data not shown).

6.3.8. Structural characterization of HerA and CptA

Structural exploratory studies by electron microscopy were performed on HerA and CptA in collaboration with Dr. J. Castón and C. P. Mata, from the CNB in Madrid. Both proteins were overexpressed in *E. coli* BL21 using pET28b derivatives pAB135 and pAB201, respectively (Annex I) and purified by affinity chromatography (Fig. S6.5, Annex IV). For electron microscopy assay, ~75 µg of each protein were incubated at 65 ºC with ATP (1-10 mM) for 30 min in a step that was compulsory to observe the hexamerical ring-shaped structures shown in figure 6.10.

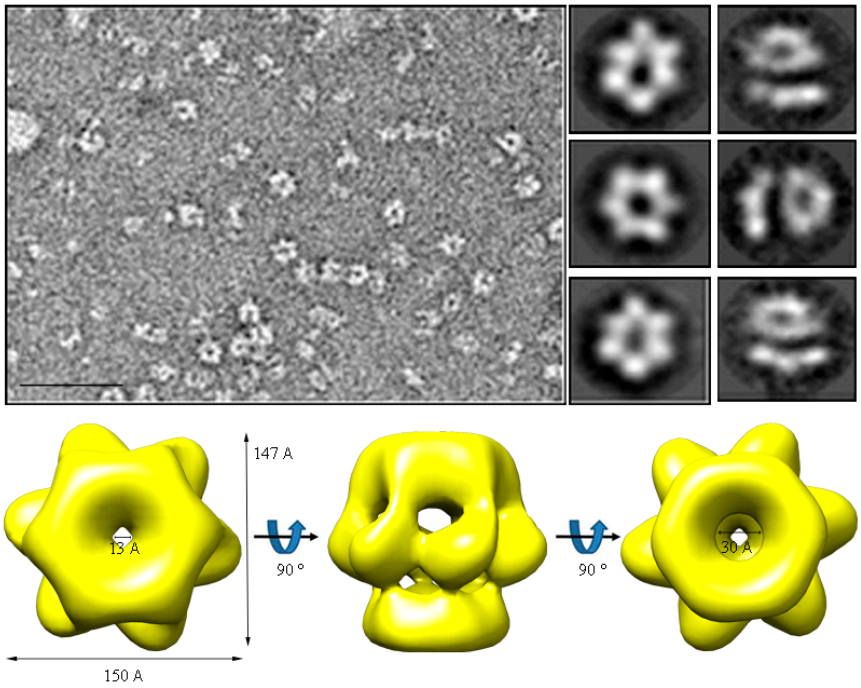
Particles of HerA preparation showed an uniform size without fragmented individuals. However, we also inspected a HerA sample purified by size exclusion and witnessed identical results as the non-size sieved ones. Structure reconstructions of HerA included 86,506 particles for the 2D models, from which 41,897 had a predicted average size of 147 Å and showed a C6 symmetry ring shape (150-95 Å) with a central pore of 13-30 Å of diameter (Fig. 6.10.A). Each monomer entailing each ring particle was, on average, 46 Å of diameter. Profile views evidenced the two modules in which the particles were conformed, explaining the variability of the sizes estimated. These particles and their related symmetries were used to calculate 3D models, which confirmed this 2-component structure of HerA. A major upper part composed by symmetric rings of 150 Å of diameter, and a bottom module depicting a rather conical structure where bottom ring sized 95 Å in diameter and vertices were smoother.

Likewise, CptA structure was reconstructed from 103,550 particles. However, in this case, larger heterogeneity among particles profile leaded the 2D models. As it can be observed in figure 6.10.B, there was a greater diversity among the ring classes and the monomer aggregation profiles, as multiple particles depicted dimer and trimmer conformations. This high variability among particles probably hampered achieving a higher resolution model. Actually, the model obtained seemed to lack the bottom module of the protein. Note that detection of these proteins during purification

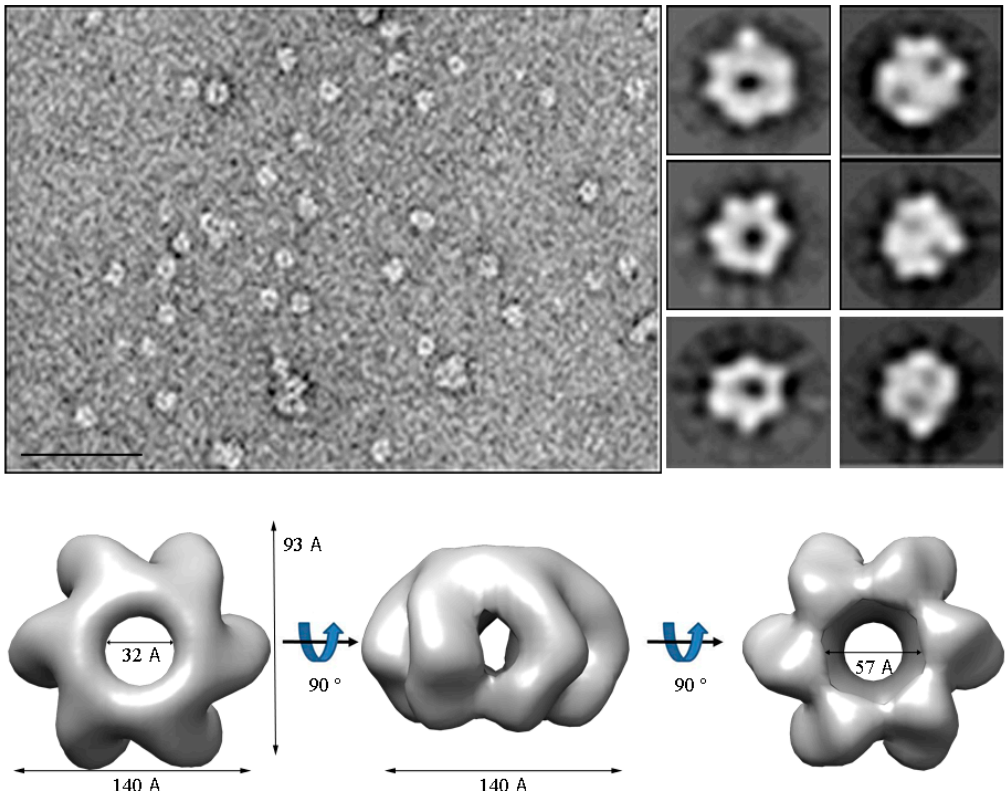
confirmed that HerA and CptA were the same size (Fig. S6.5; Annex IV), which disagreed with the model reconstructed from CptA particles, according to the sizes estimated from structural data. Regardless of this likely flexible zone of the protein, a conserved symmetric hexameric ring ~140 Å in diameter which harboured a central pore of ~32 Å on the top view enlarged to ~57 Å of the bottom view. These sizes were in the same range as those estimated for the upper ring-shape part of HerA, although embracing a potentially wider central pore. Nonetheless, the techniques employed for these structural reconstructions were too intrusive and special caution was taken for size interpretation.

Both protein models were compared among each other and to other homologues which showed high fitting affinity (Fig. S6.6; Annex IV). However, fittings of the 3D structures from homologues to HerA and CptA such as *Thermus* ATPases or YjgR hypothetical proteins in *Chloracidobacterium thermophilum* or *T. scotoductus* were not available in the PDB database or did not adequately adjusted to the resolved structures from HerA and CptA, except for the FtsK-protein from *P. aeruginosa*, and HerA and NurA proteins from *Sulfolobus solfataricus* (Fig. S6.6; Annex IV).

A



B



C

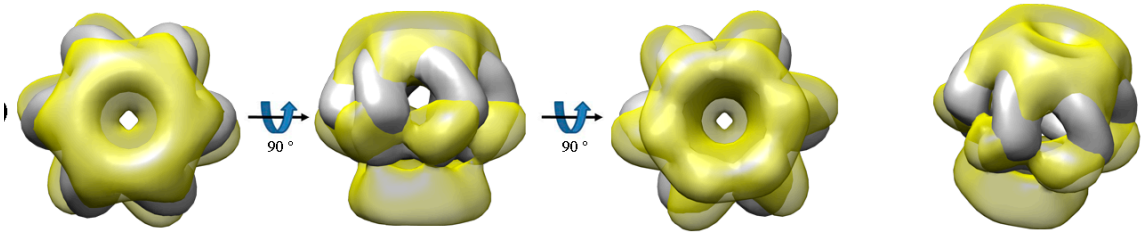


FIGURE 6.10. Protein structure reconstruction from 2D and 3D models. HerA (panel A) and CptA (panel B) proteins were purified, incubated with ATP to activate the hexameric ring conformation and adequately diluted for electron microscopy observation. Multiple high-resolution non-overlapping images were taken, and digitally processed using Xmipp software. Prototypical single particles employed for the 2D and 3D modelling were shown at the right side of both panels, over the 3D predicted structured, in which several protein views were shown and size-characterized. **C**. HerA and CptA models were matched between them.

Chapter 7

Putative role of insertion sequences in conjugation and development of a transposon-based mutagenesis system in *Thermus thermophilus*

The work enclosed in this section was partially performed at Dr. Mazel's laboratory within the "Unité plasticité du génome bactérien" at the Pasteur Institute in Paris (France).

CHAPTER 7: PUTATIVE ROLE OF INSERTION SEQUENCES IN CONJUGATION AND DEVELOPMENT OF A TRANSPOSON-BASED MUTAGENESIS SYSTEM IN *Thermus thermophilus*

7.1. SUMMARY

The presence of transposases in the cluster enclosing *cptA* encouraged the examination of their role in the potential mobility of such region. Besides, the search for active transposases was enforced by the need to develop a transposon-based tool as an alternative approach to screen mutants in conjugation, once the insertional genomic library had failed. We identified transposable elements in HB27 and HB8 strains, as well as genetic footprints of transpositional events, outlining the history of genomic rearrangements and pinpointing potentially active transposases. With this knowledge, we developed the transposon-based random mutagenesis system, enlarging *T. thermophilus* genetic toolbox.

Moreover, in the context of this thesis, the contribution of transposases to HGT was examined, as we hypothesized that the overrepresentation of Insertion Sequences (ISs), in particular in the pTT27 megaplasmid, could be contributing to the anomalies detected in conjugative transfer for certain markers. Therefore, we constructed *recA* mutants and explored the possible relevance of non-homologous integration of the transferred DNA driven by transposases. We observed that DNA was transferred between cells impaired in recombination, suggesting the existence of site-specific integration of the DNA transferred by conjugation.

7.2. BACKGROUND

After the DNA has been successfully transferred to a recipient cell (T-DNA), it is established in the host genome either by recombination or by self-replication[200]. Integration by homologous recombination is the most common way of integration and takes place when high sequence identity with resident sequences is found, resulting in sequence replacement. However, non-homologous recombination of the transferred DNA can also occur. Site-specific recombination is catalysed at specific short *att* sequence sites, mediating reversible integration when the T-DNA is flanked by two *att* sites equally oriented[173]. Transpositional recombination, catalysed by DDE motif in many transposases, involves the movement of a piece of T-DNA in the genome[24].

Presence of any of these elements is really common in many of known conjugative plasmids and ICEs and its existence either in the donor or the recipient cell, linked to the T-DNA, could drive genome rearrangements during the integration along the HGT process.

High level of transposon-mediated genome rearrangements seem to be a common trait in thermal environments[143,188], probably contributing to the extraordinary plasticity observed among the native organisms from these habitats. In fact, several studies have reported ISs and active transposases in *Thermus* species[12,28,91,98,202] including the *in vivo* transposition via conjugation of the naturally-occurring Tn*916* in *T. aquaticus*[176], although never reproduced, and the *ISTth7* transposition in *T. thermophilus* HB8[91]. Indeed, hypothetically, *Thermus* pan-genome pool might be enriched in these transposable elements in order to facilitate their lateral transfer and subsequent integration, in addition to putative phage contribution to transposon abundance through transduction. Thus, gene shuffling mediated by transpositional jumps could be a relevant contributor of HGT processes.

During the search for putative elements involved in conjugation, we found numerous truncated transposases, specially abundant in the pTT27 megaplasmid. Indeed, we identified what seemed to be IS scars (direct repeats, DR, or inverted repeats, IR) as well as partial or full transposase-like ORFs in the proximity of some of the transferred genes employed in HGT experiments, which could be involved in increased frequencies of non-homologous integration. Besides, two *ISTth7* transposases located in the megaplasmid have been reported as active. As it will commented in next chapter, in the HB27^{EC} strain, the gene encoding the Argonaute protein in HB27 is interrupted by the insertion of an *ISTth7* impeding the surveillance driven by this protein. A second case are the transposable elements flanking the HB27 denitrification island, which has been reported to be HGT-transferable[5,157]. Therefore, there were chances that the actual process we were witnessing was a combination of homologous recombination with/or a transpositional jump of the transferred targeted gene. However, lack of sufficient knowledge on *Thermus* mobile genetic elements restrained any realistic estimation on site-specific recombination of the T-DNA. For this reason, a first step was to investigate the incidence of these transposable elements. This information provides priceless information to untie the genetic history of genomic dynamism observed in *Thermus* spp which, ultimately, will facilitate comprehension on thermophilic genetic plasticity.

In summary, knowledge acquired on the insertion sequences found in *Thermus* genomes would enable the identification of potentially active transposases and the

examination of its horizontal gene transfer upon recombination-deficient strains. Besides, we would gain insights on the mobility of the *cptA* cluster, in addition to a better understanding on the pace of conjugation frequencies of several markers. Finally, this knowledge was also crucial for designing the mutagenesis transposon-based tool in order to avoid recombination with those found in the genome, hence, jeopardizing the random trait of the genetic tool.

7.3. RESULTS

7.3.1. Incidence of Insertion Sequences (ISs) in *T. thermophilus*

In order to explore the IS acquisition and how transposition events could have contributed on shaping *T. thermophilus* genome, we performed an *in silico* study on HB8 and HB27 genomes. We compiled the published material about transposable elements of *T. thermophilus* and examined the available information at the ISfinder database, which encompasses a descriptive characterization of the ISs found in HB8 genome and in pTT27 megaplasmid of HB27 strain, in addition to specialised websites such as ACLAME, MITE-Hunter, RepSeek, Repeat Finder or Repeat Masker. Despite the abundance of online resources, the majority of web browsers seemed useless as their algorithms are based on sequence homology of DNA repeats which cannot be applied in *T. thermophilus'* genomes due to the high G+C content and highly enriched in repetitions the genomes are by default. Therefore, all data obtained was compared and manually verified.

In accordance to the information compiled, HB27 genome contains 53 putative IS elements[98], from which 25 are complete IS copies (further details are enclosed in a file attached to the digital version of this dissertation). However, these data have been revisited [28,114] raising figures of the identified IS elements, recognising 9 complete copies of ISs just in the pTT27 megaplasmid of both strains. Indeed, if we consider partial IS, the amount increases up to 64 and 59 just in the pTT27 megaplasmid from the HB27 and HB8 strains, respectively. The number of partial copies found in the chromosome was lesser, around 20 copies. Thus, highest relative abundance of IS sequences (both complete and partial) was found in pTT27 megaplasmids, representing nearly a 7 % of its sequence, in agreement with the abundant presence of ISs in extremophilic habitats[21].

Despite having similar G+C content than the chromosome, the diversity of these ISs is large, classified in 10 different types of ISs, according to a combination of

features, including similarities of their ends (terminal IRs), length of the direct repeats (DR), marked identities or similarities in the transposases, including the conservation of the DDE motif (common acidic triad presumed to be part of the active site of the transposase), among others (Table 7.1).

TABLE 7.1. Typologies of ISs in *Thermus*. The main characteristics describing each type of IS is exposed in the table, including the subgroup and family to which they belong, the size of the direct repeats (DR) and the amount of ORFs comprising the insertion sequence. Besides, the re-count of ISs, both complete and incomplete, found in HB27 and HB8 genomes (HB27/HB8).

IS *Thermus*	Family	Sub-group	Typical size-range (bp)	DR(bp)	Nº ORF	Complete IS	Partial IS
ISTth1	IS3	IS150	1200-1600	3-4	2	1/0	1/1
ISTth2	IS4	IS10	1200-1350	9	1	2/0	1/0
ISTth3	IS1634	-	1500-2000	5-6	1	4/0	0/0
ISTth4&5	IS256	-	1200-1500	8-9	1	1/8	1/0
ISTth6	IS630	-	1000-1400	2	1-2	5/2	4/0
ISTth7	IS5	ISH1	900-1150	8	1	6/7	1/3
ISTth8	IS701	-	1400-1550	4	1	4/1	1/1
IS1000A/B	IS110	-	1136-1558	2	1	2/2	39/61

As it can be observed, *Thermus* specific IS types are numerically named from *ISTth1* to *ISTth8*, while two types of ISs belonging to the IS110 family, named IS1000A and its derivative IS1000B, are found too in *Thermus*' genomes, in particular, partial copies in the pTT27 megaplasmids (Fig. S7.1; Annex IV). Highest variability was observed among the 3' end of the ISs and among the target site duplication, the unique hallmark for each DNA transposon. In this way, 6 types of *Thermus* ISs were classified as DDE transposases (a common type of Prokaryotic ISs), whereas the remaining 5 belonged to families in which transposition worked otherwise. Actually, except from IS1000B and *ISTth1*, all typologies were found in both strains (Fig. 7.1)

In terms of number of copies, *ISTth7* was the most abundant. Besides, numerous copies of *ISTth4* and *ISTth5* were found in both strains, in particular in the HB8. These ISs, from the IS256 family, shared a 65 % of sequence identity, which explains their joint analysis. Many copies of these two types were concentrated in certain highly variable regions of the genome, thus, being potentially active (Fig. 7.2). Conversely, *ISTth1*, belonging to the IS3 family, was the least abundant with only one full copy found in HB27 genome.

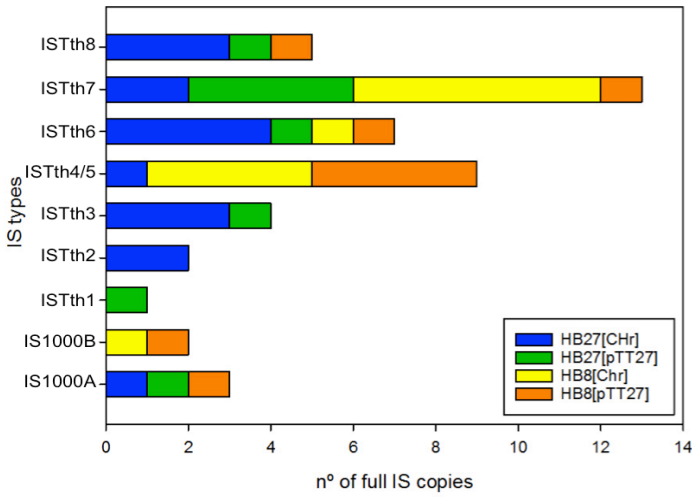


FIGURE 7.1. Abundance and classification of complete ISs found in *T. thermophilus* HB27 and HB8 strains. Following a colour code, the amount of ISs from each genome compartment (chromosome and pTT27 megaplasmid) from HB27 and HB8 strains is distributed according to its type. High similarity among *ISTth4* and *ISTth5* hindered the distinction of some of the ISs found, forcing its unification in the same category.

Moreover, figure 7.2 also portrayed IS-enriched regions found in the chromosome in contrast to the rather scattered IS distribution within the pTT27 megaplasmids. It was frequent to find interdigitation of various intact or partial IS copies in these regions. This incidence which could be rather fortuitous or may reflect scars of consecutive but isolated transposition events probably as a result of gene flow[19,126]. This plasmid-associated diversity agreed with the higher presence of transposable elements within plasmids[61,83] and with the fact that pTT27 harbours the highest plasticity with low synteny among strains[28]. Some of these IS scars were identified close to predicted genetic islands by Island Viewer, such as the denitrification island[157].

In summary, there is an over-representation of a diverse assortment of ISs in *T. thermophilus* genomes, many of them partial, concentrated in the pTT27 megaplasmids, which harbour the largest differences among HB27 and HB8. Major number of IS copies was detected of *ISTth4* in HB8 pTT27 whereas *ISTth7* was the most abundant in HB27. This relative abundance of intact IS copies, together with their location within the genome, may initially suggest its active state and also could constitute regions susceptible to homologous recombination intra- and inter-molecules (chromosome *vs.* megaplasmids).

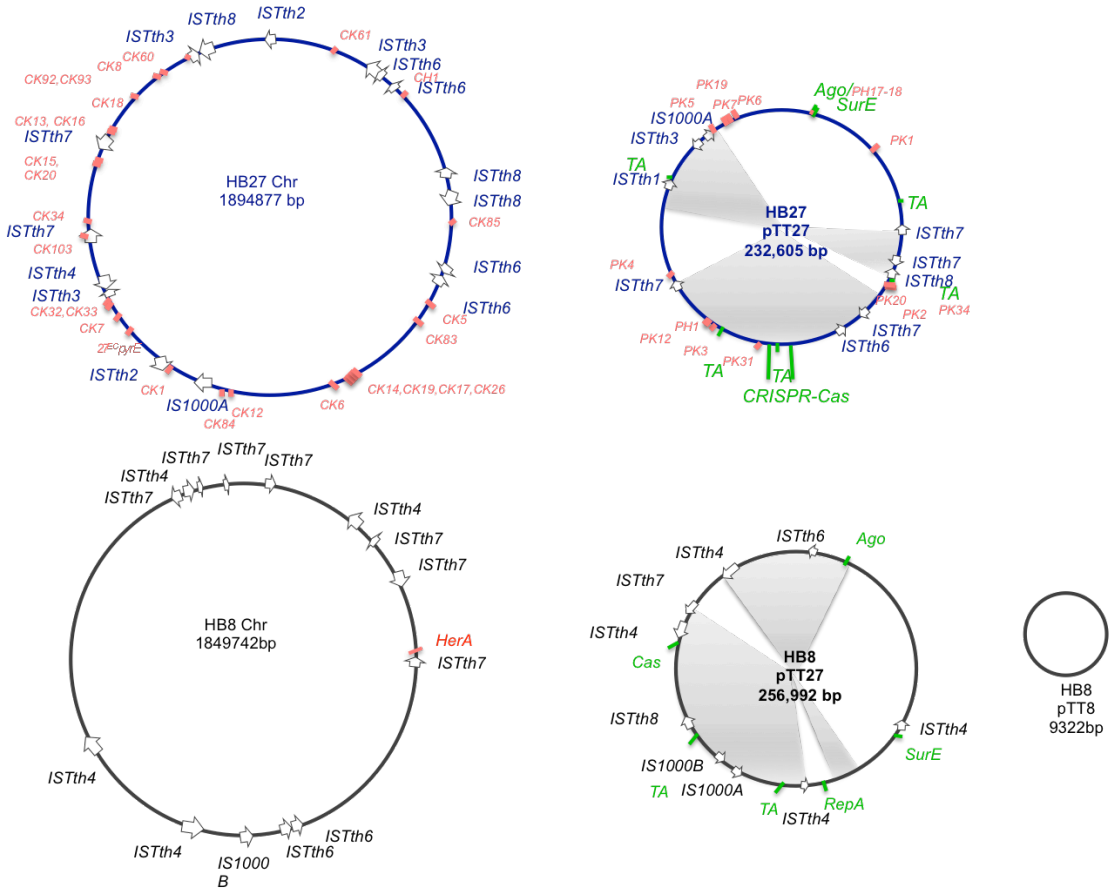


FIGURE 7.2. Distribution of complete ISs in *T. thermophilus* HB27 and HB8 genomes. ISs are indicated with an arrow oriented according to their transcription. In red, loci used in this dissertation; in green, relevant loci which may contribute to the genetic plasticity of the pTT27 megaplasmids; indicated in grey the most variable regions, showing lower synteny among *Thermus* strains.

Based on the criteria of IS abundance, localization and intrinsic features of their transpositional mechanisms (*i.e.*, target sequence size, IS sequence size, sequence conservation, etc.), we found complete copies of *ISTth4* and *ISTh7* in HB8 and HB27 (enclosed in the digital material), respectively, within regions showing higher transposable activity harbouring putative mobile genetic elements, t-RNAs, laterally transferred traits, lower G+C content, etc., especially within the megaplasmids. This suggested their potential current active state within each genome, for instance, by its insertion in the *ago* gene in HB27^{EC} strain[189], as aforementioned.

Therefore, these two types of ISs, may have a relevant role in the genome rearrangements conferring the high plasticity showcasing *T. thermophilus.* However, their consideration for the transposon tool had to be re-thought according to the high risks of recombination between copies, thus masking actual transposition as well as jeopardizing the randomness of the genetic tool, compulsory for a mutagenesis system. All in all, these results suggest that it is likely that transpositional jumps and transposition-mediated integrations can occur in *T. thermophilus*, in addition to

insertions mediated by homologous recombination, which will be studied in the following section.

7.3.2. Role of RecA in HGT in *T. thermophilus*

To test whether the laterally transferred DNA in *Thermus* cells could be integrated into the host genome by non-homologous recombination, we employed freshly isolated mutants in Δ*recA::kat* HB27 and NARI in transformation and conjugation tests.

Transformation assays involving the wild type CK1 and its *recA* derivative CK34 (HB27 Δ*recA::kat*) with either replicative plasmid pMHPnqosGFP or genomic DNA isolated from the strain CH1 (HB27 *ΔTTC0313::hyg*) were carried out. Results showed that competence in *ΔrecA* derivative was still operative, although frequencies per ng of DNA were limited by the decrease of viable cells (>2.5 fold). When these figures were expressed by viable cells and ng of DNA employed (Fig. 7.3) we found that transformation with the replicative plasmid did not rely on the presence of RecA (*p*-value: 0.052), whereas there was a significant decrease of transformation with isogenic genomic DNA (gCH1; ~12 folds; *p*-value< 0.001). These data show: i) that RecA is not involved in transformation and ii) that a significant fraction (1 out of 12) of linear isogenic DNA that enters the cell integrates in a RecA-independent manner. If we consider the number of molecules transferred, genomic transformation is more than 10 times more efficient than plasmid one, even in a *recA*$^-$ background. These results agree with previous results from our laboratory, where HB8 *ΔrecA* cells could acquire the plasmid pMK18 despite the loss of viability of the *recA* mutants[17].

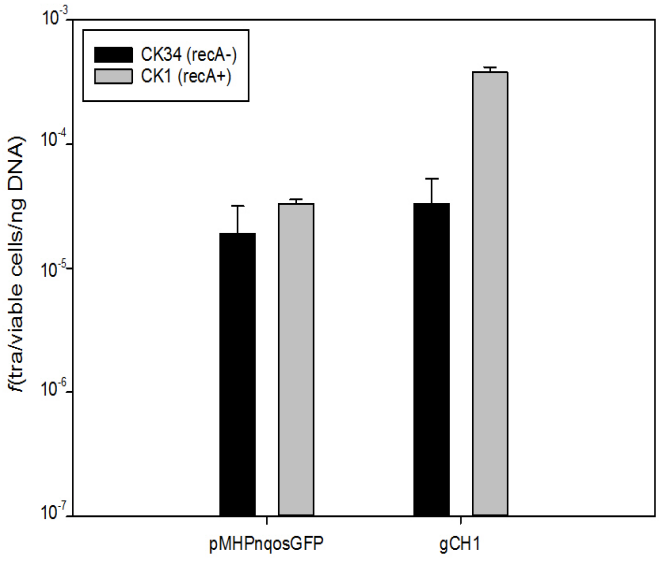


FIGURE 7.3. *T. thermophilus* can acquire DNA via natural competence independently from RecA. Transformation frequencies are expressed as number of CFU grown in TBHK (transformants) per viable cell grown in TBK per ng of DNA employed as template. Bars represent the average from 5 experiments.

Mating conjugative experiments involving the wild type CK1 and this *recA* mutant (CK34) mixed with wild type cells harbouring the pMHPnqosGFP (PpH2) or CH1 (HB27 *ΔTTC0313::hyg*) showed no difference between *recA*$^+$ and *recA*$^-$ transfer frequencies (*p*-value: 0.509), but the significance of this data was hampered by the bidirectional nature of conjugation in *Thermus* (data not shown).

RecA-independent conjugation DNA transfer of the NCE

To examine whether the CK34 (*ΔrecA::kat*) strain could work as recipient in conjugation, we performed new matings between CK34 and the pTT27- megaplasmid labelled PH1 strain (*recA*$^+$, *ΔTTP146::hyg*) and examined the transconjugants. Conjugation frequencies were similar to those observed among *recA*$^+$ strains (~6.8*10^{-4}) and ~40 % of transconjugants showed *recA*$^-$ phenotypes according to UV-sensitivity tests. These data supported that conjugative integration of the *hyg* cassette could occur in the absence of RecA.

We further tested the impact of the absence of RecA in both mates. We isolated a *recA::kat* mutant derivative of the nitrate respiring strain NARI strain (CK35) and mated it with CKH1, a *recA::kat* derivative of the aerobe HB27 that also harbored a *hyg* marker in the pTT27 plasmid (*TTP146::hyg*) (Fig. 7.4). Selection by anaerobic growth with nitrate in presence of both antibiotics (Chapter 3) gave rise to colonies that were further tested for their ability to reduce nitrate to nitrite. Results revealed that 7 out of 24 clones were positive. Parallel control experiments in which the strain CK35 (NARI) was mated with a *recA*$^+$ HB27 strain labelled with *hyg* and *kan* (CKH2), the frequency of nitrate respiring transconjugants was more than doubled (15/24).

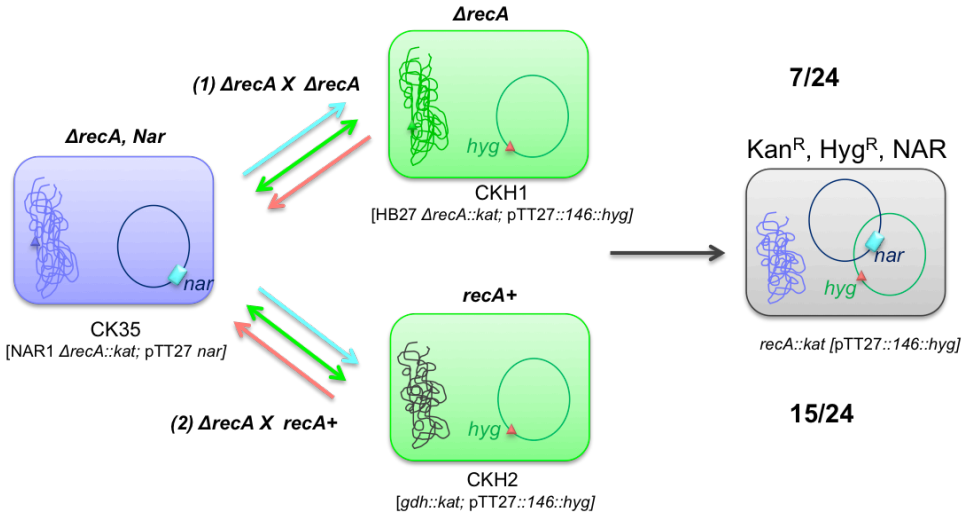
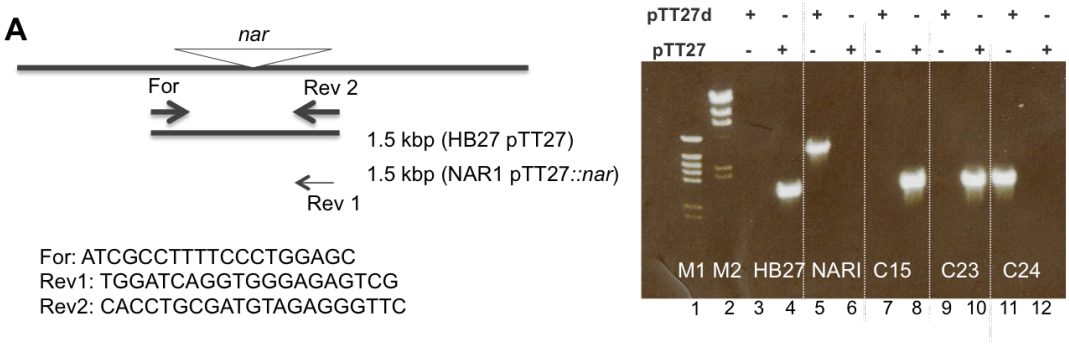


FIGURE 7.4. Transfer of the NCE in *recA*$^-$ mutants. Schematic representation of the experiments, where transfer of the NCE is represented in blue (*nar*); the transfer of the *hph5* cassette, conferring *hyg* resistance in red and the pTT27 is represented in green.

These results supported first, that RecA is also not required for conjugation and second, that integration of the NCE can take place through a *recA*$^{-}$ independent pathway.

Nevertheless, as the *nar* genes are part of the pTT27 megaplasmid, there were chances that the whole pTT27 was transferred and established in the recipient strain as a self-replicative element. In order to check this, we designed two specific PCR reactions, one to detect the NARI megaplasmid and the other to detect the pTT27 region of HB27 where the *nar* genes would have been integrated by homologous recombination, as defined in previous works of our laboratory (HB27c strain, NCE localized within *TTP097* and *TTP099* loci of pTT27 HB27)[6,157]. As shown in figure 7.5, this PCR experiment was based on a common primer (For) that annealed in both megaplasmids and two alternative reverse primers (Rev1 and Rev2), each one specific for each megaplasmid; the first one annealing immediately upstream *TTP099* and a second one annealing to a *nar*-cluster associated gene present in NARI, and hence specific for NARI (Fig. S7.2; Annex IV).



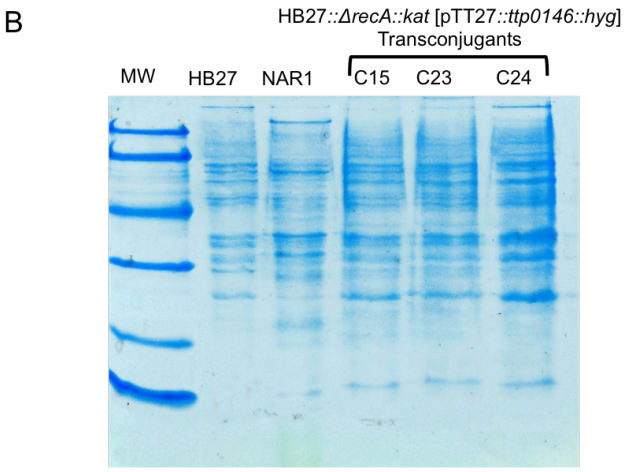


FIGURE 7.5. PCR-based experiment to detect NCE integration site in transconjugants. A. Schematic representation of the PCRs performed, including the sequence of the primers employed and at the right side, the electrophoresis of the PCRs products, where parental strains CK35 (NAR I) and CKH1 (HB27) were indicated in addition to 3 of the transconjugants tested. (C15, C23 and C24). Molecular weight markers are referred as M1 and M2. **B.** Membrane pattern profile of parental and transconjugants represented in A section. A molecular weight marker (MW) was also included.

Membrane protein patterns of the transconjugants confirmed the HB27 strain as the recipient. Amplification results indicated that in all the transconjugants analysed, there was a single PCR product, indicating the presence of a single megaplasmid species (Fig. 7.5). Moreover, in the nitrate-respiring transconjugants obtained in a

ΔrecA background, a 1.5 Kbp amplicon was found (lanes 8, 10 and 11), supporting that integration of the NCE took place in other regions of the pTT27 or the chromosome than those described for the HB27c or HB27d strains, thus, demonstrating that, as expected, integration took place by a RecA independent way, likely through the activation of ISs.

7.3.3. Development of a transposon mutagenesis tool as an alternative strategy to generate random insertional mutants in *T. thermophilus*

The construction of a thermostable transposon-based system would enable performing random mutagenesis in *Thermus* strains. Development of this genetic tool was supervised by Dr. Didier Mazel. Design of this novel synthetic transposon tool was scheduled following methodologies largely supported on the literature based on antibiotic resistant suicide constructs[67,166]. In contrast to the recently published HB27 transposon mutagenesis system by Carr *et al*[36], no previous *in vitro* transposition was arranged, whereas we expected a natural jump of the transposon.

Hence, the proposed prototype involved a non-replicative vector in *Thermus* which carried the *kat* gene for the thermostable Kan resistance. Flanking the antibiotic resistance we placed the corresponding sequences recognizable by the transposase selected, which had to be included in the vector in *trans* under the control of a promoter from that controlling the *kat* gene. A second non thermostable antibiotic resistance was also to select the plasmid in *E. coli*.

Similar to the pUT vectors[99], the Cm^R pSW23 plasmid[69] was selected as the template vector. This small cloning vector (1582 bp) replicates in *E. coli Π1* (Table 3.1, Chapter 3), that carry the protein required for recognition of its OriV. The *kat* gene used was under the control of the promoter and terminator of *rpmB* gene (*TTHA0042*) of *T. thermophilus* HB8.

Regarding the selection of transposases, we initially considered choosing one of the copies from those IS types identified as potentially active within HB27 and HB8 genomes. However, we were concerned about the high chances of recombination, would jeopardize randomness of the transposition. This point could be somehow solved by the employment of synonymous genes [33] but the high complexity of the transpositional pattern followed by our best *ISTth* candidates, discarded its employment of *Thermus* ISs. Thus, we performed a search for suitable transposases according to their size, transpositional pattern, G+C content and growth temperature of

its origin strain. Transposases *ISBStI* and *ISNth2*, from *Bacillus stearothermophilus T-6* and *Natranaerobius thermophilus JW/NM-WN-LF*, respectively, and its corresponding sequences of insertion were finally selected. *ISBstI* belongs to the IS481 family and *ISNth2* is an ISLre2 type transposase. These two candidates were the most suitable for the purpose of the project and fulfilled several of the criteria aforementioned, in particular a simple transpositional pattern, which would ease transposition occurrence and its subsequent detection (Table S7.1). Expression of both transposases was under the control of Pnqo promoter[44] and a Rho-independent transcription terminator was added to their 3' extreme. The resulting constructs pAB131 and pAB132, are sketched in figure 7.6 and their nucleotide sequence is enclosed within the digital material provided..

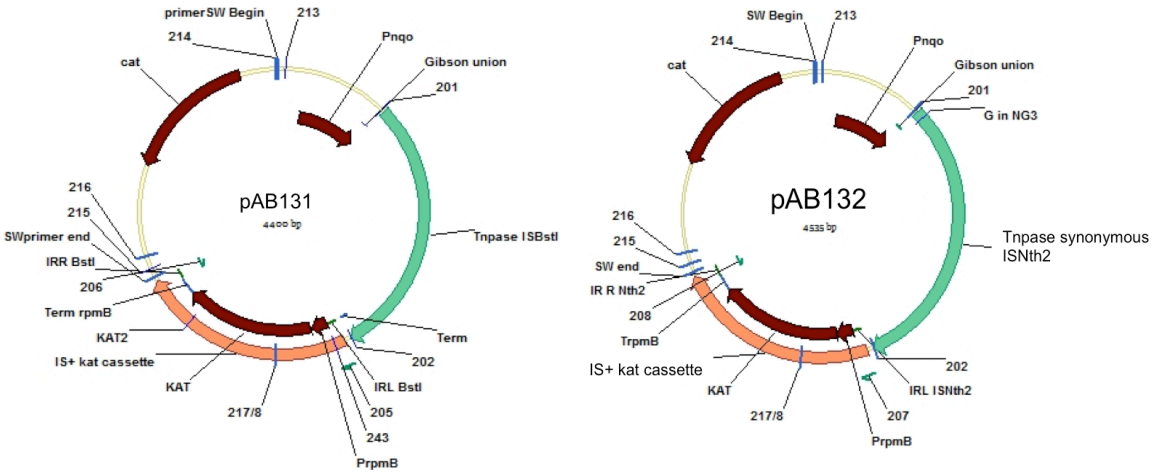


FIGURE 7.6. Transposon-based random mutagenesis system for *Thermus* strains. Maps of pAB131 and pAB132, indicating, in green, the transposase, in orange the complex of the *kat* cassette (promoter and kat gene highlighted in brown) flanked by the left and right inverted repeats, indicated as IRR or IRL, depending on the orientation. In brown, the genes encoding the Chloramphenicol and Kanamycin resistance (*cat* and *kat*) and the promoters (Pnqo and PrpmB). Terminators and Gibson joint sites are also indicated, in addition to the oligonucleotides employed, which can be checked in Annex II.

Functional transpositional excision and random insertion of the transposases, compulsory for this transposon-based mutagenesis tool, was verified by the analysis of insertion sites in 10 HB27 isolates obtained after trasnfomration with these plasmids. We detected that the transposases had inserted in different sites of the HB27 genome (Fig. 7.7). Once verified the accurate performance of the constructs, they were used to generate a library of mutants that was stored. As natural transformation provided low yields of Kan^R transformants with both constructs, we used electroporation instead. Currently, we have accrued 1,729 clones for *ISBstI* and 1,109 for *ISNth2*. Future work through massive sequencing methods will define a map of insertion of each of these transposases.

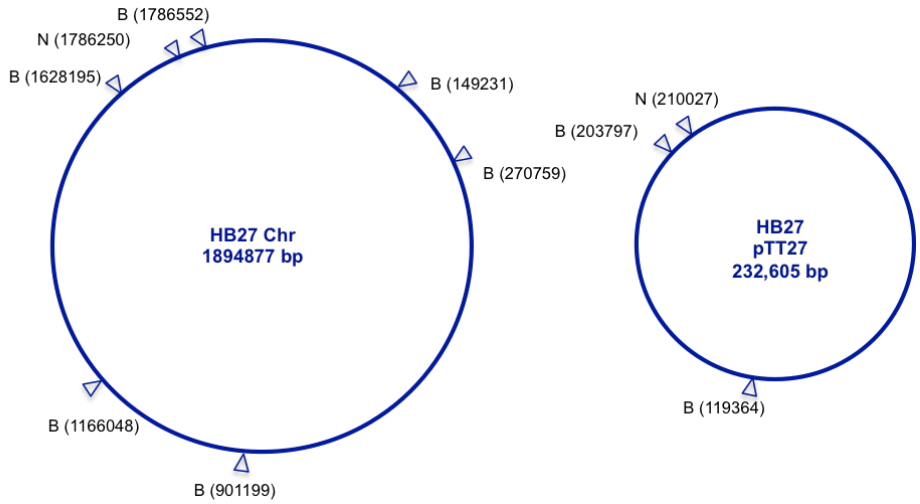


FIGURE 7.7. Validation of the transposon-based random mutagenesis system for *Thermus* strains. Insertion sites of the transposon in 10 different clones are indicated according to their position in the sequence of the chromosome and the megaplasmid of the HB27 strain.

Chapter 8

Discriminative interference mediated by the Argonaute protein in *thermus thermophilus* horizontal gene transfer

Parts of this section have been published in Blesa, A.; César, C. E.; Averhoff, B. and Berenguer, J. 2015. "Non-canonical cell-to-cell DNA transfer in *Thermus* spp. is insensitive to Argonaute-mediated interference" published in *Journal of Bacteriology 197 (1), 138-146.*

and Blesa, A. and Berenguer, J. 2015. "Vesicle-protected extracellular DNA contributes to horizontal gene transfer in *Thermus spp*". *International Microbiology*. In press

CHAPTER 8: DISCRIMINATIVE INTERFERENCE MEDIATED BY THE ARGONAUTE PROTEIN IN *Thermus thermophilus* HGT

8.1. SUMMARY

In contrast to its eukaryotic counterparts, the Argonaute protein encoded in *T. thermophilus* (TtAgo) protects the bacterium against DNA acquired by transformation through DNA-DNA interference. TtAgo presence eliminates nine out of ten DNA molecules entering the cell, regardless the origin of the DNA or whether it is embedded within extracellular vesicles. In this chapter, we analyse the role of TtAgo in conjugation and in the physiology of *T. thermophilus*. We show that, not only there is an evident role of TtAgo in the physiology of the bacterial cells, but also, that, in contrast to transformation, this DNA-DNA interference mechanism is not elicited when the DNA is acquired by conjugation. A putative candidate for ssDNA guide supplier to mediate the target interference is also examined.

8.2. BACKGROUND

In Eukaryotes, Argonaute (eAgo) proteins lead the RNA-RNA interference systems that protect the organisms from viruses and transposons, and that is also involved in the regulation of gene expression of certain genes. Besides, Archaea (32 %) and bacteria (9 %) encode eAgo homologues known as prokaryotic or pAgo[190]. Among the best characterized pAgos is that encoded in *T. thermophilus* (TtAgo), which was used as 3D model for human Ago 2[191]. Crystallization of TtAgo showed that, as long as a ssDNA guide is provided, this protein performs cuts in the complementary strand in what was described as the first DNA-DNA interference mechanism. Furthermore, *in vitro* studies evidenced that despite TtAgo could target both DNA and RNA, it functions better as ssDNA-DNA interference mechanism, cleaving efficiently at A+T rich sequences of the complementary DNA strand[178]. *In vivo* studies performed by Swarts *et al*[189] have shown that this cleavage affects the DNA incoming by transformation, increasing over an order of magnitude the number of transformants when TtAgo is deleted and suggesting an active role of TtAgo in cell defence, comparable to that reported for the CRISPR-Cas systems.

TtAgo is encoded in the pTT27 megaplasmid as a monocistronic gene *TTP026* in HB27 and *TTHB068* in HB8 strains. In other *Thermus* strains, even in those with almost negligible natural competence, such as SG0.5, TtAgo is well conserved. Sequence comparison revealed a high degree of identity among all TtAgo homologues (Fig. S8.1; Annex IV) (from 99.7 % in HB8 to 43.1 % in *T. scotoductus* SA-01). In all cases, TtAgo was associated to a megaplasmid (*T. thermophilus* strains *JL18, SG0.5, HB27, HB8, Thermus* spp. Isolate 2.9 and *NMX2.A1*) or associated to regions of ancestral megaplasmid which had been integrated into the chromosome (*T. scotoductus* SA 01, and *Thermus* spp. CCB US3 UF1, for instance). The genetic context of HB27 TtAgo encoding genes, akin to the aforementioned homologues, suggests its expression as an independent transcriptional unit, which regulation remains yet unknown (Fig. 8.1).

In order to untie any additional functions of TtAgo and its putative role in defence against DNA transferred by conjugation, we carried out a thorough study of TtAgo mutants.

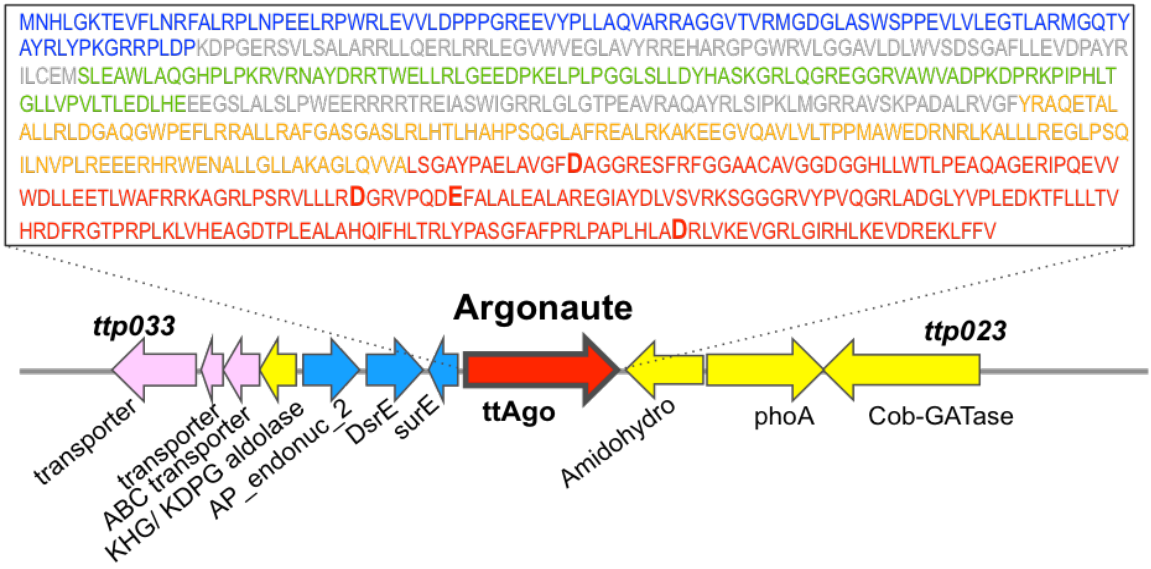


FIGURE 8.1. Genomic context of the gene encoding the TtAgo protein in *T. thermophilus* HB27.. The colour code for the arrows, representing *ORFs*, is the following: in red *TTP026* gene, coding for TtAgo; in yellow, genes coding for proteins involved in metabolic pathways; in blue, those with putative DNA binding motifs and in pink an ABC transport system. Sizes pictured are proportional to genomic dimensions. The aminoacid sequence of HB27 TtAgo is also included; in blue the N-terminal domain, in green the PAZ domain, in yellow the MID domain and in red the catalytic PIWI domain. Grey represents the two linkers flanking the PAZ domain. Catalytic residues are highlighted in bold.

8.3. RESULTS

8.3.1.Physiological characterization of TtAgo mutants

As a first approach we characterized the phenotype of the *Δago* strain and compared it to its HB27wt parental. We performed a battery of tests with both strains and included in every assay the HB27^{EC}strain, an spontaneous mutant in TtAgo which carries an *ISTth7* insertion, as well as the MD158 strain[37], which lacks half of the pTT27 megaplasmid including the *ago* gene.

We assayed temperature dependence of growth and, in all cases (60, 65 and 70 ºC), similar trends were obtained, featured by an slightly faster growth of *Δago* strain. Overall, generation times at 65 ºC were 48.17 ± 2.51 and 47.03 ± 2.63 min for the HB27wt and a *Δago,* respectively. Cell yields at the end of the growth were also similar (in average, 1.08 ± 0.29*10^{10}). HB27^{EC} behaved similar to *Δago* whereas MD158 showed lower growth rates at all points checked.

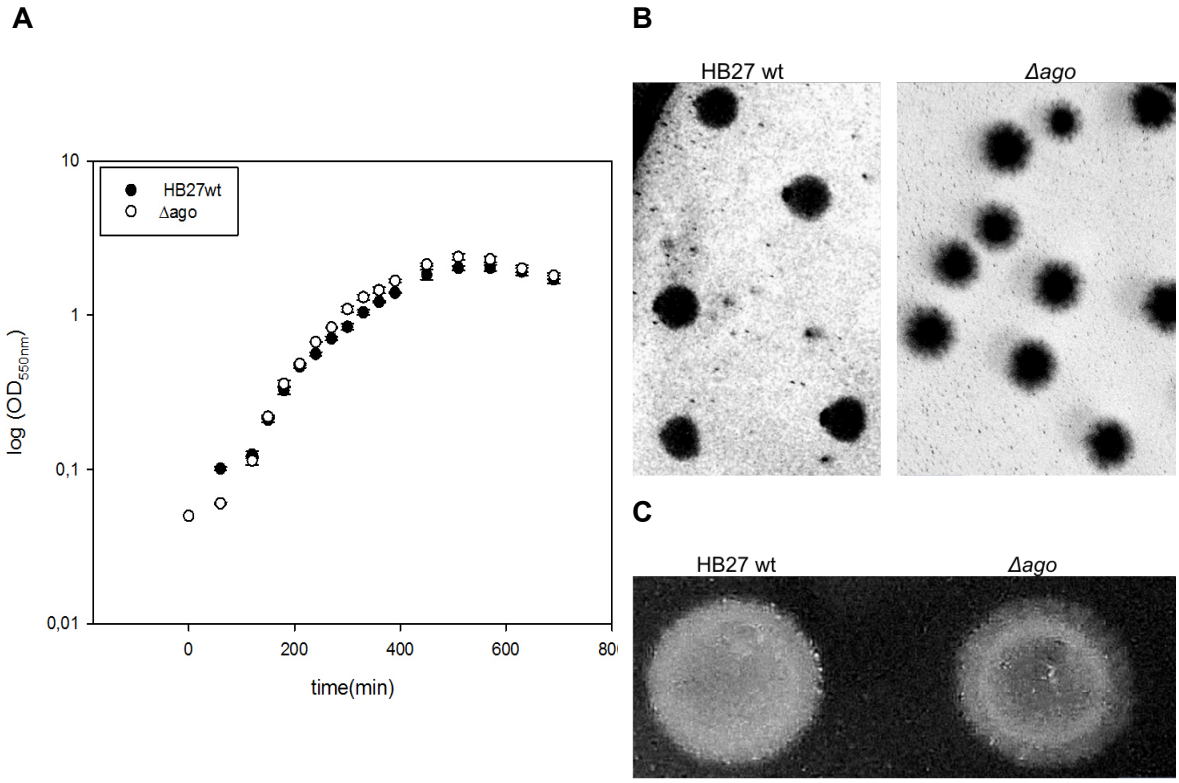


FIGURE 8.2. Growth of Ago variants at 65 ºC. A. Growth curves represented as the logarithmic scale of OD measured at 550nm against time (measured in minutes) since start (OD_{550}=0.02) of HB27wt (*ago*$^+$) in black dots and *Δ ago* (*ago*$^-$), in white. Error bars represented the deviation of the mean of three independent samples. **B.** Isolated CFU of HB27wt (left side) and *Δ ago* (right side) strains grown in TB plates. **C.** TB plates spotted with 10µl of each saturated derivative indicated above the image.

Besides, morphological differences were evident among colonies of the wild type and the *Δago* derivative on TB plates at 65 ºC (Fig. 8.2.B and C). Either the single colonies (Fig. 8.2.B) or the spotted inoculum of the *Δago* strain (Fig. 8.2.C) showed more irregular margins featured by a contoured halo with rather curled borders, suggesting increased motility. No further difference on pigmentation could be observed

among both cultures, which on stationary phase were yellowish and in liquid medium did not form flocks or any other kind of cell aggregations. Likewise, HB27^{EC} strain did not show any deviation to the aforementioned profile for the *Δago* strain. In contrast, MD158 was white, evidencing a significant loss of pigmentation, as expected from the lack of part of the carotenoid biosynthesis pathway encoded in the missing section of the megaplasmid[37]. Contrarily, biofilm formation measurements of the *Δago* strain were nearly two times higher than those estimated for the HB27wt strain (Fig. 8.3.A). The capacity of biofilm formation of derivatives HB27^{EC} and MD158 was also higher than that of the wild type strain (not shown). Analogously, adherence and twiching areas shown by the *Δago* strain were visually larger (Fig. 8.3.B). The spreading zone of the *Δago* was larger and less uniform than that shown by the wild type strain, which was around 1.48 ± 0.22 cm. Sub-agar-adhesion assays, based on growth of cells attached to the Petri dish beneath the agar layer, portrayed similar results, evidencing larger adhesion areas for the *Δago* strain. These results may infer that TtAgo enhance type IV-pili twiching-motility. In fact, double mutants in *Δago* and genes encoding pilins (*pilF, pilT, pilQ, pilA4*) were defective in twiching motility and adherence, supporting a role for TtAgo in piliation. However, there were *ago*$^{-}$ mutants impaired in certain pilins too such as CK27 (*pilT*) and CK7 (*pilT2*), that were hyperpiliated but defective in twiching motility, as reported by Salzer *et al*[168].

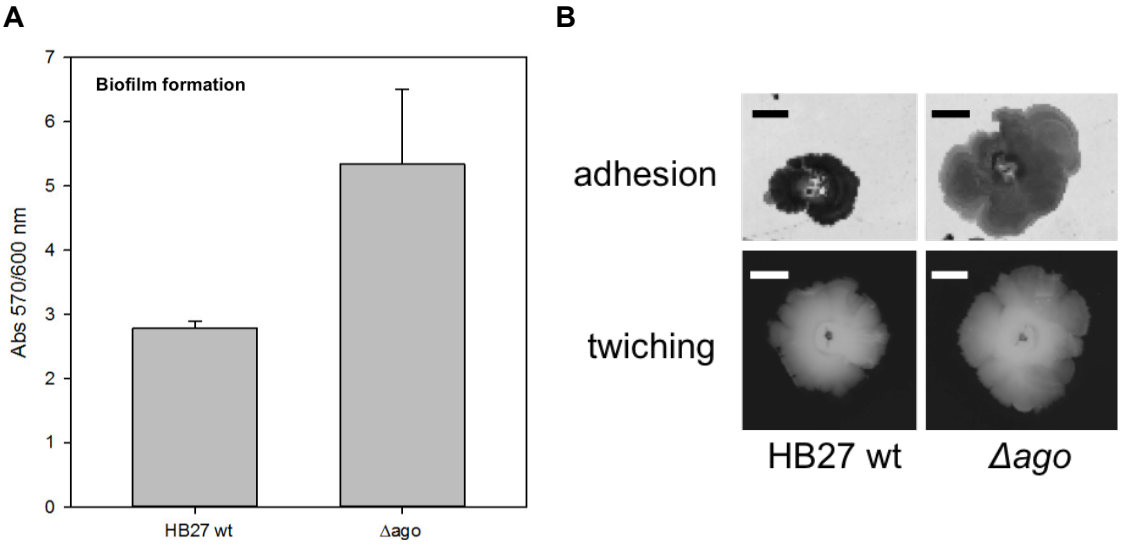


FIGURE 8.3. Effects of *ago* impairment on motility and cell aggregation in *T. thermophilus* HB27. **A.** Biofilm formation of HB27wt (*ago*$^{+}$) and *Δago* of saturated cultures was estimated as the ratio between the absorbance of crystal violet dye measured at 570 nm by the OD_{600} of each culture. Values represent the mean of 3 different experiments, each one entailing 3 replicates. **B.** Adhesion (upper panels) and twiching-motility (lower panels) of cells of HB27wt (left-side panels) and *Δago* (right-side panels). For twiching-motility assays, agar and cells were dyed with Coomasie-blue and then rinsed. The spreading zone corresponded to the non-coloured agar region. For adhesion tests, agar was removed and the petri dish was washed and adhering cells were stained with Coomassie blue. These are representative images of triplicates. Bars represent 0.5 cm.

In contrast, the response of TtAgo under environmental stresses, such as UV light irradiation or abrupt raise of growth temperature (from 60 ºC to 79.2 ºC), did not reveal significant differences. No differences could be observed when exposing cells to UV light (60 $J \cdot m^{-2}$, 15 cm distance) on qualitative and quantitative UV-resistance tests, where ratios of survival yields of UV-exposed cells to those non-radiated counterparts (UV-radiated CFU/viable CFU), were in the same range ($1.07 \pm 0.47 * 10^{-1}$ and $0.95 \pm 0.31 * 10^{-1}$, for HB27wt and *Δago* respectively). Likewise, similar survival yields from HB27wt and *Δago* strains towards the sudden temperature raise, examined at three time lapses (Fig. S8.2, *p*-value: 0.44). Growth curves of HB27wt and *Δago* strains at 79.2 ºC followed similar trends and yields did not significantly vary between them (data not shown).

In summary, we observed that deletion of TtAgo resulted in an increase in biofilm formation and in twiching motility and a slight enhancement of growth. Perhaps, these differences together with likely further physiological effects yet unknown envisage a farther role of TtAgo in cell physiology and support its genomic conservation through evolution in *Thermus* spp.

8.3.2. Role of TtAgo in Horizontal Gene Transfer

To elucidate the physiological role of TtAgo in DNA exchange, in particular in conjugation, we performed a battery of HGT assays with several *ago*$^-$ mutants from HB27.

8.3.2.1. DNA-DNA TtAgo interference in transformation is not dependent on the nature or structure of the incoming DNA

To examine whether different TtAgo derivatives interfered with different types of DNA taken up by natural competence, we performed transformation assays using a variety of templates: plasmids, isogenic genomic DNA, ssDNA and DNA associated to EVs (Fig. 8.4).

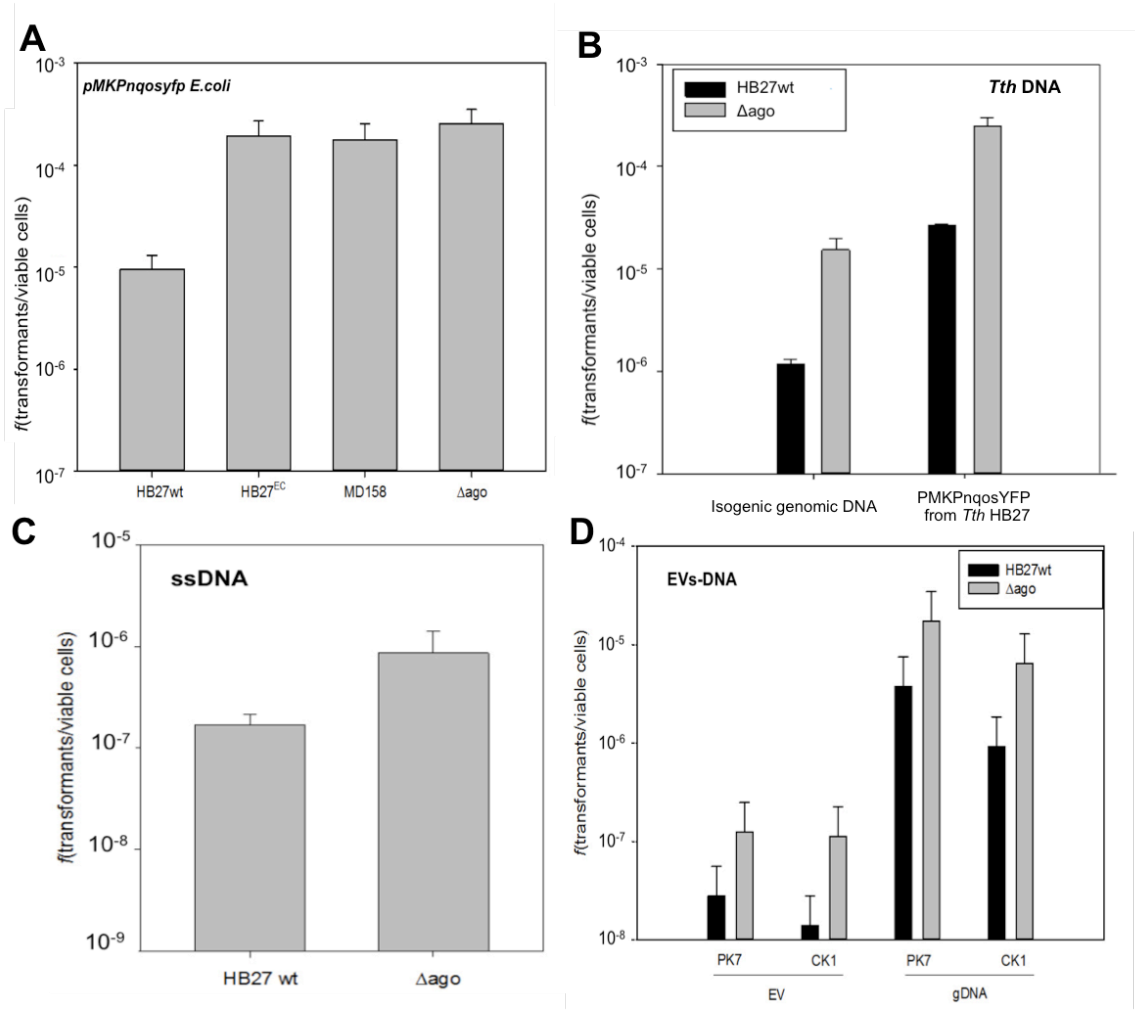


FIGURE 8.4. TtAgo interferes with DNA taken up by natural competence. Transfer frequencies in the indicated strain were expressed as the CFU of transformants per the CFU of viable cells. **A**. HB27wt (*ago*$^+$) and *ago*$^-$ strains *Δ ago,* HB27^{EC} and MD158 were transformed with 200 ng of plasmid pMKPnqosYFP purified from *E. coli.* **B.** HB27wt (*ago*$^+$, black bars) and *Δ ago* (*ago*$^-$, grey bars) were transformed with either 20 ng of isogenic genomic DNA from CK1 strain and 200 ng of plasmid pMKPnqosYFP purified from HB27wt. **C**. HB27wt (*ago*$^+$, black bars) and *Δ ago* (*ago*$^-$, grey bars) were transformed with ssDNA obtained by linearization and boiling of 2.5 μg of plasmid pAB66. **D.** Transfer frequencies of HB27wt (*ago*$^+$, black bars) and *Δ ago* (*ago*$^-$, grey bars) with EVs fraction (EV) from the CK1 and PK7 strains containing 500 ng of eDNA. Transformation controls with 20 ng of genomic DNA from same strains were carried out in parallel (gDNA). Error bars correspond to the mean standard deviation (n=3).

As it can be observed in figure 8.4.A, transformation with plasmid pMKPnqosYFP purified from *E. coli* was ~10 fold higher in the *ago*$^-$ derivatives compared to the wild type strain (*p*-value< 0.01). No differences could be detected among the different *ago*$^-$ strains (*p*-value: 0.08). However, it had also been reported that preferential cleavage sites of TtAgo were A+T-rich[178]. In order to elucidate if this preference was accomplished *in vivo* we performed transformation assays with genomic DNA from isogenic strains. As revealed in figure 8.4.B, isogenic DNA also elicited TtAgo interference (*p*-value< 0.01). Moreover, when plasmid pMKPnqosYFP purified from *T. thermophilus* was employed, the difference between *ago*$^-$ strains and

the wild type was also, at least, one order of magnitude. In consequence, these results show that TtAgo's activation was independent from the restriction-modification systems, and also from the topology and G+C content of the incoming DNA. When isogenic ssDNA (figure 8.4.C) was used, a similar cleavage trend with around ten times higher transformation efficiency of the *ago*$^-$ derivative compared to the wild type one (*p*-value< 0.01) was observed. Indeed, TtAgo DNA-DNA interference was also active towards isogenic DNA associated to extracellular vesicles isolated from HB27 derivatives (figure 8.4.D).

Discriminative interference shown by TtAgo may be supported by results from electroporation assays, where both *Δago* and *ago*$^+$ electro-competent cells showed similar acquisition rates of a kanamycin marker (Fig. 8.5). This artificial entrance of the DNA into the cell did not elicit TtAgo interference (*p*-value: 0.33).

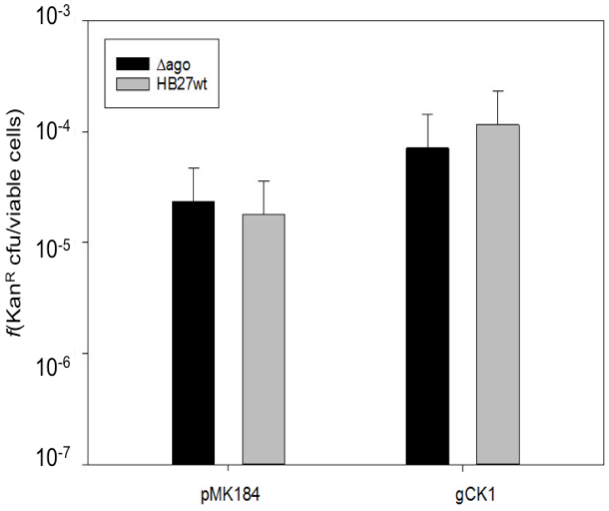


FIGURE 8.5. Electroporation does not elicit TtAgo DNA-DNA interference. HB27wt (*ago*$^+$, grey bars) and *Δ ago* (*ago*$^-$, black bars) electro-competent cells were electroporated with either 20 ng of genomic DNA from strain CK1 and 100 ng of plasmid pMK184 and grown for 4 hours at 60 ºC before plating. Transfer rates correspond to the mean from 3 independent experiments, calculated as the number of Kan^R CFU by the total viable cells of each strain. Error bars represent the mean standard deviation. No statistically significant difference was observed in any case between *ago*$^+$ and *ago*$^-$ rates (*p*-value: 0.333).

In summary, TtAgo DNA-DNA interference system was active not only against low G+C plasmids as published, but also against isogenic high G+C DNA, both as dsDNA and ssDNA forms, and regardless it is embedded or not within vesicles when these molecules are incorporated to the cell through the competence apparatus.

8.3.2.2. Conjugation does not elicit DNA-DNA TtAgo interference

As we have seen in Chapter 5, *Thermus* spp. can horizontally transfer DNA by a conjugation-like process in which the competence machinery is required in the recipient strain. We hypothesized that this dependence of conjugation on the transformation apparatus of the recipient cell could elicit TtAgo interference.

To examine this, we performed mating assays involving wild type (*ago*$^+$) strains labelled in the chromosome with Kan^R (CK23) or Hyg^R (CH6) and their *Δago*

counterparts (CK22 and CH5, respectively). Transformation tests were performed as controls. As shown in figure 8.6, no significant differences among the transfer efficiencies were found between matings of *Δago* and those of *ago*$^+$ strain, whereas the expected decrease of an order of magnitude in transformation efficiency was observed in the control experiments with *ago*$^+$ strain. Same results were obtained irrespective of the position of the antibiotic marker in the mates (Fig. S8.3; Annex IV).

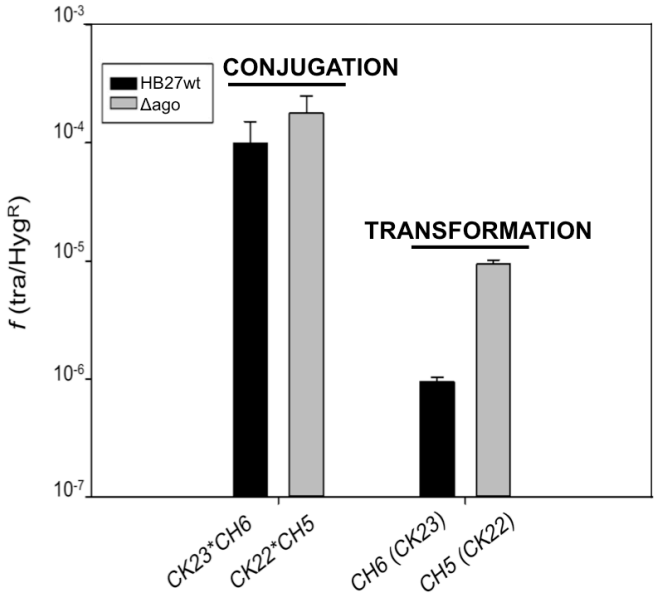


FIGURE 8.6. TtAgo discriminates between DNA acquired by natural competence from that incoming by conjugation. Transfer frequencies were obtained after crossing equal cell amounts of Δ*ago* (CK22*CH5) strains as well as *ago*$^+$ strains labelled in the same genes (CK23* CH6). Parallel transformation of CH5 and CH6 strains with 10 ng of genomic DNA extracted from the respective counterparts (CK22 and CK23), were carried out as controls. Frequencies represent an average of 9 independent experiments. Error bars correspond to the mean standard deviation. Differences of conjugation transfer frequencies between Δ*ago* and *ago*$^+$ crosses were statistically non-significant (*p*-value: 0. 50; n=9). However, there are significant statistical differences among transformation frequencies (*p*-value< 0.01).

Moreover, *Δago* derivatives followed the general conjugative attributes defined in Chapter 5, such as the preference of transfer for pTT27- associated genes or the requirement of competence for at least one of the mates which would work as putative receptor (Fig. S8.4; Annex IV). Likewise, similar results were obtained for *ago*$^-$ derivatives HB27^{EC} and MD158 derivatives.

All in all, these results indicated that TtAgo could distinguish between the DNA that enters the cell by natural competence from that donated through contact by a neighbour cell, even though in both processes the transformation apparatus is employed.

8.3.3. Expression and localization of TtAgo

8.3.3.1. TtAgo is expressed constitutively

Availability of α- TtAgo antiserum obtained by our group allowed us to follow its expression *in vivo*. First, TtAgo's expression along HB27wt growth cycle revealed no significant change in its expression (Fig. 8.7). In addition, we did not detect any significant change upon a battery of tests that were performed including heat shock, cold shock, UV-light, as well as presence of *Δago* cells or DNA (data not shown).

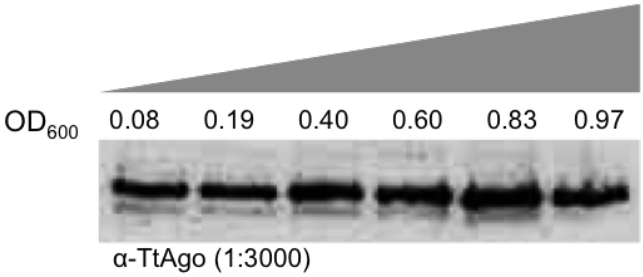


FIGURE 8.7. TtAgo expression along growth cycle. Immunoblotting of equal cells amounts collected at different points of HB27wt growth cycle, expressed as the OD measured at 600nm. α - TtAgo dilution employed was 1:3000.

8.3.3.2. Localization of TtAgo

To localize TtAgo within the cells, we followed two strategies; the first one based on cell fractionation and the second, involving fusions to thermostable fluorescent proteins.

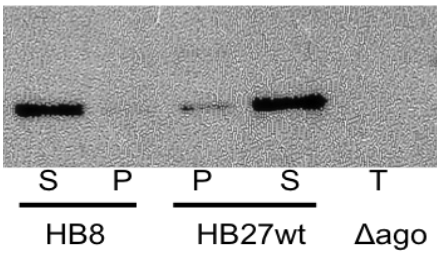


FIGURE 8.8. Cell fractionation immunoblot for TtAgo subcellular detection. Cell extracts from different saturated cultures of *T. thermophilus* strains (HB8, HB27wt, *Δ ago*). Samples were processed and separated into independent fractions: total (T), soluble (S) and insoluble (P).

As it can be observed in figure 8.8, most of the TtAgo protein in HB27 and HB8 strains was found in the soluble fraction, but with a faint band detected in the insoluble fractions (lane P) in both strains, which suggests the ability of TtAgo to interact with the membrane. As expected, no detection was achieved on the *Δago* strain, used as negative control.

In the experiments with fusions to sGFP and sYFP we used a monocopy insertion into the chromosome (pAB214 and pAB221, respectively). As shown in figure 8.9, no defined region of fluorescence accumulation could be deduced in any of the constructs used, which rendered an homogeneous distribution along the cells.

Functional assays of strains harbouring these fusion constructs were checked by Western-Blot (Fig. S8.5; Annex IV) and in complementation assays, where the fusion complemented the wild type phenotype in transformation.

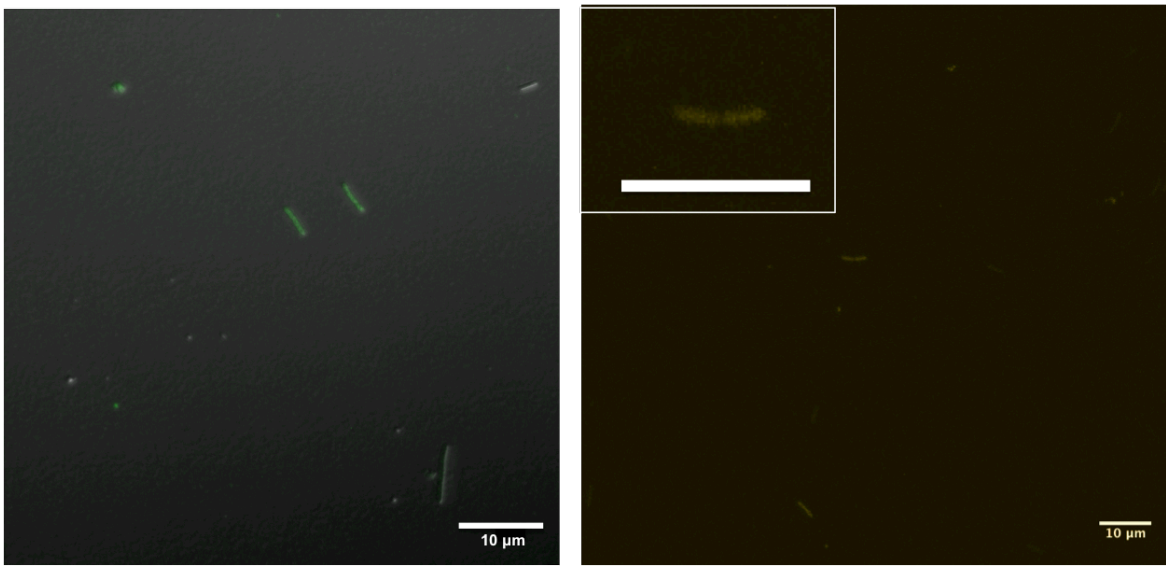


FIGURE 8.9. Fluorescent localization of TtAgo in HB27. Confocal microscopy images from four different fusion strains of *TTP026* (encoding HB27 TtAgo) to sGFP (left) and sYFP (right). Around 10^7 cells from single recombinants were fixed and laid onto ultrathin agar layers on top of confocal microscopy preparations. Scale bars are represented in the right-side bottom corners, indicating to 10 µm.

8.3.4. TtPrimPol, a putative DNA guide supplier to TtAgo

Several *in vitro* assays proved that TtAgo required the ssDNA guide to perform the DNA-DNA interference[191]. However, the origin of such guide remains unknown, prompting us to search for potential candidates for guide suppliers. We thought that the PrimPol, described by the group of Luis Blanco (CBMSO, Madrid) as capable to generate *de novo* small ssDNA, encoded in locus *TTC0656* in the HB27 strain, could potentially co-work with TtAgo. In order to test this, we constructed single and double mutants (*Δago, ΔPrimPol*) and tested them in transformation assays.

Replacement of TtPrimPol gene (*TTC0656*) by a *kat* cassette resulted in a delay of growth and a slender reduction of viable cells in mutant CK81 (~3 fold compared to CK1). This reduction of viability became abrupt (>18 fold) in the double mutant CK82 (*ΔPrimPol,* ∆*ago*). In fact, colonies of these strains showed a loss of pigmentation as well as a trend to aggregate in small flocs, in particular, the double mutant CK82. Nonetheless, these pleiotropic effects did not hinder the performance of transformation assays (Fig. 8.10). Results showed that impairment in TtPrimPol enhanced competence, but surprisingly, to levels that even surpassed that of the ∆*ago* mutants regardless of the DNA template employed. In contrast, double mutant CK82 (*ΔPrimPol,* ∆*ago*) showed a small decrease in transformation with the plasmid. Differences between transformation efficiencies were also evident when the DNA template was isogenic genomic DNA. In this case both primPol mutant (CK81) and the

double primPol and ago mutant (CK82) showed a similar increase in transformation efficiencies as that shown by the single ago mutant (CK22) compared to the wild type. These results support a role of the PrimPol in the same interference pathway of TtAgo.

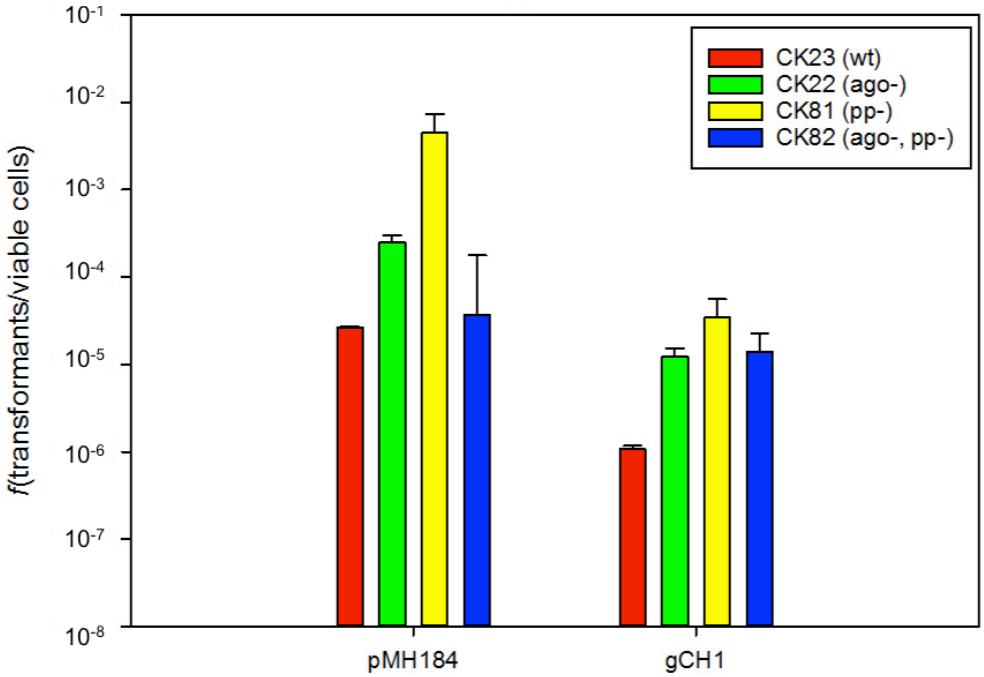


FIGURE 8.10. Competence experiments of TtPrimPol mutants. Transformation frequencies of PrimPol^{+} CK23, CK22, and their PrimPol^{-} counterparts (CK81 and CK82), were estimated as the number of transformants (Hyg^{R} + Kan^{R}) per total viable cells (Kan^{R}). Frequencies represent an average of 8 independent experiments. Error bars represent the mean standard deviation.

Chapter 9

Discussion

CHAPTER 9: DISCUSSION

Thermophilic niches sustain many organisms showing an extraordinary plasticity, consequence of the fluent horizontal flow of genetic material within and among populations[16]. Prominent genetic promiscuity requires both, versatile systems of DNA acquisition (and its further integration), and effective surveillance systems, which protect them from potentially hazardous DNA acquired in transfer. In this dissertation, we provided evidence showing that *T. thermophilus* exchanges DNA by a conjugative mechanism at least as efficient as its outstanding natural competence machinery**.** We also show that TtAgo, a thermophilic homologue of eAgo proteins favours the acquisition of foreign genes by this conjugation system as its DNA-DNA interference is only elicited when DNA is acquired by transformation, but not by conjugation. Besides, this discrimination system might have evolved to permit a high but safe promiscuous HGT, as observed in this bacteria. Equilibrium between looseness of the DNA acquired and the protection provided by the TtAgo system may contribute to the development of significant adaptative advantages, ultimately enhancing its evolutionary success. We also analysed the relevance of non-homologous recombination in conjugation and examined the distribution and functionality of the ISs in HGT, according to the identification of several potentially active transposition sites which may support their contribution to the intense genomic rearrangements featuring *Thermus* spp[28,114]. In the following subsections, the key findings from this Thesis are discussed.

Competence *versus* conjugation

Evolutionary fitness advantages acquired by many strains of *T. thermophilus* have been ascribed mainly to the active role of its highly efficient natural competence system[16]. However, there are some cues fuelling the hypothesis that the major role of the competence apparatus has actually evolved not as a mean to gain genetic diversity, but as a feeding system. First, it is active for the intake of all sources of DNA present in the milieu, independently of their origin or the degree of sequence similarity, thus, precluding in most cases its insertion by homologous recombination[51,198]. Second, in most bacteria where natural competence plays a role in diversification, the system is inducible under specific conditions, principally *via* quorum sensing, such as nutrient

limitation or DNA damaging stress[74], allowing the population either to repair DNA or to acquire new traits from other strains of the same species. In contrast, natural competence in *T. thermophilus* is constitutive and takes place along the whole growth phases of static cultures[113], where DNA serves as an excellent source of nutrients for growth in this bacterium. Third, the TtAgo protein has been shown to function as a limiting factor for the integration of extracellular DNA into the cells as it interferes at least in nine out of every ten molecules that enter the cells[189], thus, reducing the chances of gaining new genes by transformation. Last but not least, we should bear in mind that these bacteria live at temperatures at which dsDNA is not stable and most of it could be found mainly as single stranded, which, despite being internalized by the competence apparatus shows a much lesser transformation efficiency than dsDNA under laboratory conditions (Fig. 8.4). In this sense, the EVs described in Chapter 4 could certainly work in protection of the dsDNA released as a consequence of cell lysis, allowing a long-range and long-time window for transformation and for HGT. However, the efficiency of this EVs transformation process seems too low to have an actual relevance in HGT in natural hot environments, although its putative contribution in nature cannot be ruled out completely.

So, our starting point lied on the fact that *T. thermophilus* should had developed alternative, explicit, highly processive means to exchange DNA safely and efficiently, beyond transformation. Indeed, more than fifteen years ago, Ramirez-Arcos *et al*.[157] demonstrated the transfer of capability for nitrate respiration through a conjugation process among *T. thermophilus* strains. We believed that this lateral inheritance could not be just a restricted event, but that could have wider impact on *Thermus* biology and HGT. Actually, our comparison of efficiencies between transformation and conjugation using the same amount of total DNA than that contained in a "donor" strain (Fig. 5.2) clearly favours conjugation as the most likely HGT-pathway between *Thermus* strains.

The conjugation model

Basic experiments such as the paucity of effects in the presence of DNase I (Fig. 5.1), the requirement for cell-to-cell contacts revealed by the U-tube device experiments similar to those developed by Davis[65], or the need for two living partners, provided mounting evidence suggesting that this DNA transfer followed a classical conjugation model. Moreover, the more than 10-fold preference for the transfer of genes located in the pTT27 megaplasmid over those located in the chromosome

resembled the fluent flow of conjugative plasmids in classic systems, where the transfer of chromosomal genes is the consequence of the integration of the megaplasmid in the chromosome by recombination between same insertion sequences (IS) found in both genetic elements. Actually, this *Hfr*-like mechanism was the explanation that our group gave to the apparent order of transfer observed in interrupted mating experiments involving the transfer of the nitrate respiration conjugation element (NCE) from the NARI strain (formerly identified as HB8) and the HB27 strain[157].

However, the absence of homologues to main components of the classical TSS4–based conjugation systems in the megaplasmid nor in the chromosome of the conjugation proficient HB27 strain, supported that conjugation in this strain significantly diverged from standard processes. Cornerstone advocacy to this view was provided when we detected that conjugative exchange could take place between completely isogenic strains, thus, endorsing a system not subjected to any surface or plasmids-incompatibility exclusion mechanisms, typical of classical conjugation models[86,181].

A first approach to unveil this apparently unconventional conjugation-like mechanism in *Thermus* was to examine the role of the competence apparatus on it. The rationale for this was the complexity of the competence apparatus itself, formed by an amalgam of homologues to components of type IV pili (T4P) and of T2SS secretion apparatus[16]. If DNA transfer by conjugation required the competence machinery, some of the mutants in competence would compulsory be also defective in conjugation. Contrarily, we detected conjugation when one of the mates was non-competent (Fig. 5.5), thereby indicating the existence of an alternative system, independent from the competence machinery, needed for *Thermus* conjugation. However, as conjugation was not possible when both mates were non-competent, the competence apparatus was somehow involved in conjugation and required at least in one of the mates, either the “donor” or the “recipient”.

A major problem at this point was the futile identification of the “donor” and “recipient” roles among isogenic *T. thermophilus* mates, thus, hampering a proper parenthood analysis. To solve this, we had to use non-isogenic strains derived from HB27 and NARI strains which presented perfectly distinguishable membrane protein patterns. Taking advantage of this fact, we could demonstrate that the natural competence apparatus was required only in the recipient strain, but not in the donor.

These data confirmed that conjugation in *T. thermophilus* relies on the competence machinery to acquire DNA by the recipient cells, and encouraged us to hypothesize about the existence of an active *pushing* system of DNA in the donor which worked

independently but likely synchronized to the competence machinery of the recipient, ensuing the proposal of the *"push and pull"* model, sketched in figure 5.7.

Nevertheless, other alternatives were pondered. For instance, it could be that just some of the components of the competence machinery instead of all, were required for the release of DNA out from the donor cell. This would be in accordance to the proposed bi-functional role of certain transformation components[168]. Proof of this hypothesis requires matings involving all possible combinations of competence mutants, that is 16^2 matings if we consider the already published 16 competence proteins[16], which results are not available yet. Nonetheless, we could verify that the conjugative phenotype is independent from piliation, as both hyper-piliated (PilT2, PilT) and non-piiliated (PilQ, PilF) mutants were impaired in conjugation when crossed to a *ΔpilA4* derivative. Likewise, some mutants in proteins of the transformation apparatus that could uptake DNA at low frequencies (PilT2, ComEC, DprA, PilT), were still refractory to conjugation with a non-competent *ΔpilA4* strain. These data are not perfectly suited for rejecting the putative implication of components from the competence apparatus in conjugation, but decline the participation of the best characterized and likely main components of the systems in the hypothesized *pushing* process.

Initially, we approached the search for components of the DNA pushing system through the generation of single recombination mutants in a non-competent *ΔpilA4* background. Hence, we designed a large genomic library enclosing *T. thermophilus* DNA fragments in the suicidal vector pK18, where mutant candidates would be obtained by its electroporation on *ΔpilA4* cells. To our surprise, yields of Kan^R cells recovered after electroporation were significantly lower than expected, suggesting that these *ΔpilA4* cells were also refractory to acquire DNA by electroporation. Further trials on different competence-less backgrounds (*ΔpilQ, ΔpilF*) were of no avail, leading to hypothesize that during electroporation in *Thermus* spp, the DNA does not actually access to the cytoplasm but probably is temporarily confined in the periplasmic space, where it remains until the natural competence machinery transports it to the internal compartments of the cell. However, as it will be discussed later, this means of DNA acquisition does not elicit TtAgo-mediated interference. In any case, constraints upon the attainment of mutants in competence-less background compelled the search for alternatives, such as designing a transposon-based mutagenesis system that could increase the efficiency of the mutagenesis.

Despite the successful construction of this tool, as described in Chapter 7, completion of an *in silico-in vivo* screening entitled the identification of some of the

components involved in DNA donation during conjugation. Based on the presumption that *Thermus* conjugation should involve proteins of the same PFAM family as the chief enzymes orchestrating classical conjugation systems, we designed a pipeline guiding the screening. Of particular interest were ATPases and enzymes harbouring DNA-binding motifs, as we presumed that this process would require substantial energy to gear an efficient performance of the specialized DNA processors. In our preliminary bioinformatic analysis we identified different protein candidates which were further checked by mutagenesis. As shown in Chapter 6, we uncovered certain genes which inactivation produced abrupt decreases, or even the cancellation, of conjugation when combined with an impairment of the DNA gateway (*i.e.*, *pil*$^{-}$). Some of these mutants showed significant decreases in cell viability, prompting its discard. At the end, only two candidates were selected for further inspection, embodying notorious effects in conjugation with minor or no effects in competence or viability. Of course, we do not deny the potential chances of other still non-detected proteins involved in DNA donation during *Thermus* conjugation, including those affecting cell survival when mutated.

Our two *bona fid*e candidates encoded putative ATPases-Helicases. Among these two, CptA is apparently strain-specific and involved only in conjugation, as the mutant did not show further detectable phenotype alterations apart from the impairment in conjugation, whereas the well-conserved HerA protein would likely be involved in additional duties. Notwithstanding how complex the system could be, the identification of proteins affecting specifically DNA donation in conjugation confirms our hypothesis for a "*push and pull*" mechanism involved in conjugation in *T. thermophilus*.

The *Push* machinery

As commented above, HerA and CptA, the two candidates spotted, belong to the same ATPase family, which also includes the partitioning protein FtsK. However, in those strains in which both are present (*i.e.,* HB27) they are independently required in conjugation, that is, absence of any of them halts DNA *pushing*, suggesting a collaborative interaction between them during the transfer.

Beyond its functional resemblance, HerA and CptA show a significant degree of sequence identity (32 %) and similar hexameric conformations when exposed to ATP. Together with their putative ATPase capacity and nearly identical size, data strongly suggest a major role of these proteins as DNA helicases, an expected endeavour

within the context of ssDNA transfer during conjugation[90]. However, both proteins show different biochemical behaviours. While HerA is only found in the cytoplasmic fraction, is easily overproduced in *E. coli* and purified for electron microscopy analysis, CptA partially associate to the membrane fraction *in vivo*, and its expression and purification from overproducing *E. coli* cells faced several obstacles, ultimately hindering the resolution achieved under the electron microscopy (Fig. 6.10). Besides, there are also clear differences regarding the phenotypic effect of their respective absence; while CptA impairment did not carry further phenotypical effects in addition to the loss of conjugation capacity, deletion of HerA notably affected cell survival when exposed to UV-light. This UV-sensitivity reveals HerA's involvement in DNA repair processes. Actually, HerA encoding gene (*TTC0147*) is preceded in HB27 by a gene encoding a pseudogene homologue to NurA (Fig. 6.4), but by a full protein in other *T. thermophilus* strains, supporting the involvement of both proteins in a DNA repair system of double-strands breaks similar to that described in Archaea[58]. Actually, recent studies in *T. thermophilus*[109] and in *Deinococcus radiodurans*[52] on the role of the NurA-HerA proteins DNA repair and recombination, support this view. In this scenario, it could be possible that HerA helicase participated in the generation of the substrate, either ss or dsDNA for the transfer. Also a putative role in recombination could explain the decrease in transformation efficiency suffered by *herA* mutants (~33 fold).

Certainly, the role played by HerA in DNA repair should be well conserved along the genus, and its direct contribution to active pumping of DNA through the membrane seems unlikely. However, strains like HB8 which are also capable of conjugation (either donate and receive DNA), do not encode CptA, supporting the requirement of additional yet unknown components that surrogate CptA. In any case, it should be noted that HB27 shows meaningfully higher conjugative transfer efficiencies than any of the other strains studied.

CptA's singleness is in accordance to the available information which suggests its recent acquisition by insertion into the chromosome of the HB27 strain. Actually, CptA seems to be part of a 13-Kbp DNA fragment of lower G+C content surrounded by ISs that has been inserted into a conserved region of the HB8 genome (Fig. 6.5). This CptA's insertion seems to have spread among other strains as it was also detected by PCR in NARI, starred, as well, by an efficient conjugation capacity. These data, together with the presence of a putative restriction-methylation system within the same cluster suggest that this insertion could partially behave as an ICE[213].

Compilation of these data submitted that within *Thermus* conjugation, there are really efficient strains which harbour both CptA and HerA proteins, while there are

other strains, less efficient, where only one protein (*i.e.*, HerA) is present. Interestingly, the genes encoding both proteins are immediately preceded by putative NurA-like proteins, although that preceding HerA in HB27 appears as an inactive pseudogene (Figs. 6.3 and 6.4). These evidence, although yet insufficient to withstand a *bona fide* model, impelled hypothesizing that in strains harbouring the ICE-like insertion encoding CptA, this protein works energizing the DNA transport to the recipient cell, while HerA could be an helicase *de facto* required for the generation of the transferred-DNA copy in the donor cell or even of the actual strand to be transferred. Conversely, in strains devoid of CptA, other unknown protein would replace CptA's display (Fig. 9.1).

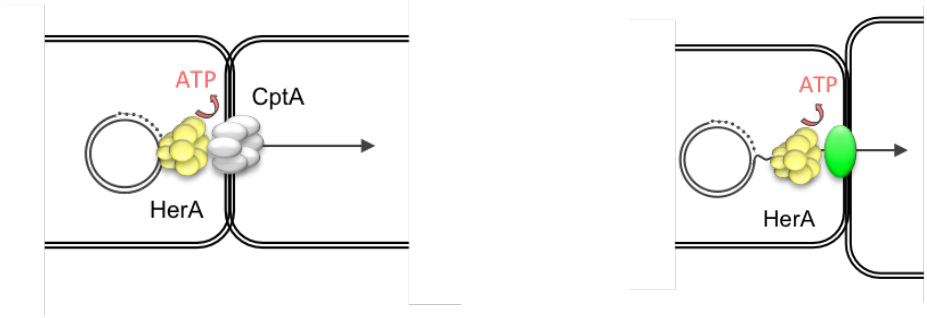


FIGURE 9.1. Sketch of the hypothetical roles of HerA and CptA in DNA donation during conjugation in *T. thermophilus*. The helicase HerA is represented in yellow whereas CptA is coloured in grey. The unknown component(s) aiding HerA when CptA is absent is pictured in green.

The predicted role of CptA in DNA transfer to the recipient cell is endorsed by its resemblance to *Streptomyces* TraB protein[196,197,205]. In these bacteria, TraB, an ATPase belonging to FtsK-SpoEIII family and homologue to *Thermus* CptA, is the unique responsible for dsDNA transfer through the mycelial septa, which, aided by a large multiprotein complex, is also driving the spread of the newly acquired DNA from mycelial apical region to other neighbouring strains. Similar to *Streptomyces* conjugative model, CptA shows a polar localization within the cell, principally in one of the poles according to the confocal microscopy analysis of single recombinant derivatives where CptA is fused to thermostable fluorophores (Fig. 6.9). In this sense, the collaborative actuation of CptA and HerA in the process of DNA transfer seems to be reinforced by the polar localization of HerA, similar to that observed for CptA (Fig. 6.9). Therefore, our hypothesis lies on the fact that conjugative DNA transfer occurs at the pole of the donor cell. Regarding the recipient cell, data provided by Dr. Averhoff indicate that the competence machinery, required for the postulated "second step" driving *Thermus* conjugation (where it is involved in the capture of DNA capture its pulling into the recipient cell) is uniformly distributed along the cell's surface[15,29], insinuating, thus, the absence of specific preferential sites to set contact with the donor

cell. However, more recent data suggest that PilQ and pili composing the competence apparatus are concentrated at the cell poles[169,88], which, if pilins would foster a role in conjugative transfer similar to that exhibited in natural competence, would picture an equivalent scenario as the one witnessed in *Streptomyces*, where DNA exchange is restricted to the poles.

A crucial point of this two-step model proposal, which has also relevant concomitant implications in the interference phenomenon mediated by the TtAgo protein, is related to the nature of the DNA transferred. Then, it is questioned whether the DNA is transferred as single-strand, alike classical models widespread among Proteobacteria and Gram-positive bacteria, or, in contrast, DNA is transferred as dsDNA, resembling *Streptomyces* TraB model. According to the structural data obtained by electron microscopy, it is feasible that CptA could afford the passage of double-stranded molecules through its central pore within the hexamer. However, according to the disparity of sizes among the structures solved of CptA and HerA, it seems evident that the structure solved for CptA lacks part of the protein, probably missing both upper and lower domains, which could also alter the actual size of the central pore observed. Up to this point, just high-resolution structures obtained by electron cryo-microscopy would empower obtaining truly reliable data regarding the CptA hexameric inner pore diameter. Therefore, with the current available data, we cannot favour any of the two possibilities considered.

Nevertheless, it seems unlikely that, in the proposed model framework, the single presence of CptA suffices to grant the successful transport of the DNA across the complex cell envelope of two *Thermus* cells. These cell walls are composed by, in addition to the cytoplasmic membrane, a rather thick (~30 nm) composite peptidoglycan layer covalently bound to secondary cell wall polymer (SCWP) to which the SlpA protein attaches, enabling the anchoring of an outer membrane[2,39-41]. In this sense, we strongly believe that ancillary components yet to be characterized would be required to smooth DNA's transport through these multilayer structures.

In this sense, there are multiple conserved genes encoding T2SS-like components among *Thermus* spp which would be involved in conjugation. For example, deletion of the putative *pulE* gene in HB27*,* fetched an actual decrease of conjugative transfer (Table 6.2). Besides, some of proteins involved in competence could be also participating in DNA release from the donor cell.

The origin of transference

Most of the well characterized conjugation systems enclose one or few short DNA sequences explicitly recognizable by DNA-binding proteins. These sequences are encoded together with the conjugative apparatus and regarded as origins of transfer or OriT(s). These sequences set the starting point of the DNA conjugative transfer by a covalent binding of the 5' end of the OriT to the relaxase, which also shows a nicking activity. Actually, it is the relaxases' recognition by the T4SS responsible for the transfer of DNA what triggers its propulsion through the T4SS channel, dragging the DNA attached to it[32]. Generally, this process is energized by the ATPases localized at the base of the conjugation apparatus, working as actual macromolecular motors. In classical models, the fact that OriT is a *cis*-acting sequence discloses an actual order in the transfer of the genes, starting with the closest locus to the 5' end to which is bound to the relaxase, until the one placed at the remotest point, corresponding to the free 3' end, generated as a result of the DNA cleavage.

In *T. thermophilus* the results obtained from the inserted markers at different loci within the pTT27 megaplasmid did not reveal a clear order in their transfer. However, the pace of transfer of chromosomal markers showed a different trend, pinpointing some regions as more frequently transferred than others (Fig. 5.3). Occurrence of these preferential zones would suggest the existence of close OriT-like sequences. Currently, further specific information is not readily available, but, we can hypothesize about putative OriT candidates. Actually, the CptA cluster encodes an ORF (*TTC1877*) 99 % identical in sequence to *Tth111II* from *T. thermophilus 111* (T. Oshima) a type IIGS restriction enzyme. This endonuclease is defined within a family of *Thermus* restriction endonucleases which are featured by a large size, an incomplete cleavage of the substrate DNA, and the fusion in a single polypeptide of both the nuclease and the methyltransferase activities[218]. *Tth111II* recognises CAA(G/A)CA sequences, and cleaves preferentially the leading strand at N11/N9, downstream the recognition sequence[217]. In HB27 genome, the CAA(G/A)CA sequence appears 53 times in pTT27 along the megaplasmid, heterogeneously distributed. Conversely, the chromosome encloses this sequence at 383 sites, but its distribution was not so scattered, finding several of them condensed at given regions. In fact, these hotspot areas of *Tth111II* recognition sites fairly coincide with those featured by higher conjugative transfer frequencies (Fig. 9.2). Also, in terms of abundance, target sites for this enzyme were proportionally more concentrated in the megaplasmid (more than one order of

magnitude difference), which could reasonably explain the transfer preference for megaplasmid-associated genes.

In this context, the efficiency of the methylation system could be chief actor in preferential transfer, exposing certain areas, reachable for cleavage, while limiting the access to others. In fact, within the same insertion as the putative type IIGS RE and immediately downstream CptA, we identified an ORF (*TTC1880*) encoding a putative methylase reported as active[10].. Actually, Furuta *et al*[85] characterized this specific cluster as a "mobile" restriction-modification system confirming its lateral insertion into HB8 chromosome, as we predicted. Probably, the integrase and the transposase flanking CptA may have promoted the mobility of this region, thus giving this insertion an ICE-like nature.

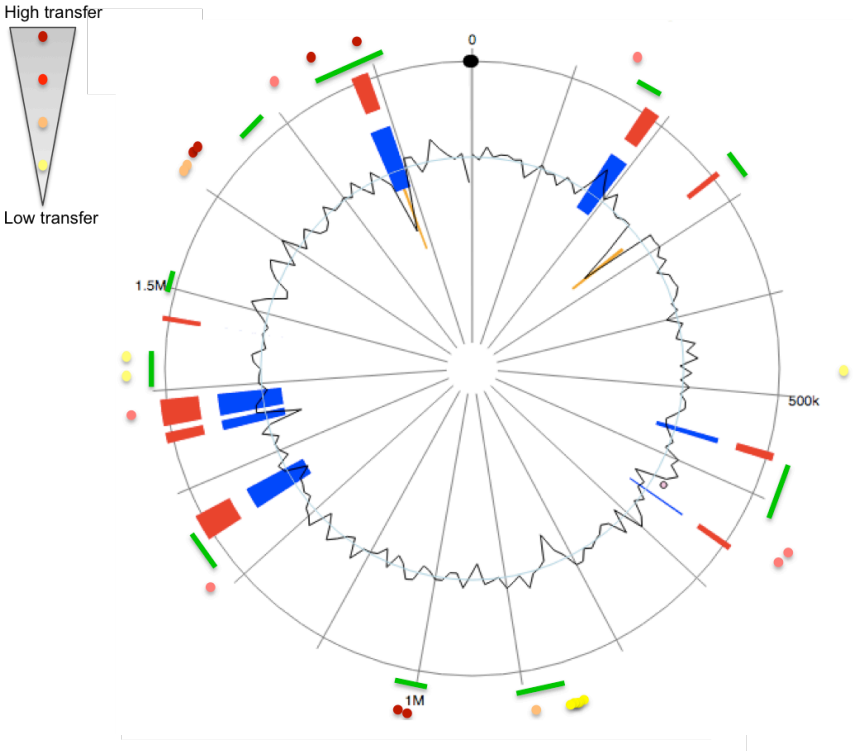


FIGURE 9.2. Distribution of *Tth111II* target sites in HB27 chromosome. Genetic map of HB27 chromosome showing the variations of GC content (grey zipped lined), potential island paths (in red) according to IslandPath DIMOB (in blue), SIGI-HMM (in orange), taken from Island Viewer. Pink dot near the GC line represents an homologue to *macB* gene (related to antibiotic resistance). Out from the circle, green lines represent the major hotspots where *Tth111II* target sites are concentrated and in a colour-degraded scale of dots, conjugative transfer efficiency of different markers.

However, an alternative or additional explanation for the preference for megaplasmid associated genes shown by the conjugation was the sporadic and likely reversible integration of the megaplasmid by recombination between any of the many IS that both elements contain, as it is described for the F plasmid *Hfr* in *E. coli* strains. Therefore, the frequencies observed for plasmid markers would be resultants of the unidirectional transfer of the plasmid marker to the chromosomally labelled recipient

mate, and only a small fringe number of transconjugants actually correspond to plasmid labelled recipients that have received the chromosomal marker. At present, we cannot distinghish between these two mechanisms.

TtAgo as a barrier for HGT

During this research, we confirmed that presence of TtAgo constitutes a barrier against incoming plasmidic DNA taken up by natural competence, as formerly published[189]. However, in that study they highlighted that *in vitro*, the interference was mainly directed towards A+T-rich sites found in low G+C content plasmids. Conversely, our experiments demonstrate that *in vivo,* the barrier interposed by TtAgo stands active against any DNA template, even if it is isogenic linear DNA either double- or single-stranded, thereby identical to the chromosomal DNA of the recipient cell. This surprising lack of specificity compels the existence of a discrimination system, as it seems obvious that TtAgo does not cleave its own genome. This discrimination system must be completely different from that identified in CRISPR/Cas schemes, where self-spacers cleavage by *Cas* is eluded by the inclusion of additional bases (PAM) found in the target but not in the spacer[194]. Therefore, if the discriminative trait does not prevail on the DNA sequence itself, differential features may abide on (i) the structure of the incoming DNA (ssDNA, for instance), (ii) spatial linkage between TtAgo protein and the DNA pulling machinery of *Thermus* conjugosome, (iii) any TtAgo activation phenomenon triggered by the uptake of DNA.

Importantly, we have also verified that when the same exact isogenic DNA which elicits the TtAgo-mediated DNA-DNA interference is transferred by conjugation, presence or absence of this protein does not affect the incoming DNA, thus, repressing, or at least not activating the surveillance defence system headed by TtAgo. If the DNA was to be transferred in a double-stranded conformation, as guessed in the former section, the no-activation response of TtAgo interference would become apparent as it would evidence the incapacity of TtAgo to act on high G+C dsDNA, hence, against its own genome. However, we have demonstrated that in conjugation processes the DNA enters the recipient cell through the same translocator apparatus as the one employed for natural competence, thereby structurally incompatible with the acceptance of both ssDNA and dsDNA. In this respect, it seems very likely that both systems lead ssDNA in the cell, fact which would not impede the exit of the DNA from the donor cell in a dsDNA conformation, also employed in transformation events.

If we undertake the former arguments explaining the lack of TtAgo's interference against its own genome, the actual distinction between DNA interference or not mediated by the TtAgo protein does not rely on the nature or the type of the transferred DNA, but on the incoming via by which the DNA enters the recipient cell. In this regard, it should be reminded that similar to the response (actually lack of it) observed in conjugation, DNA taken up by electroporation also overtakes TtAgo's surveillance barrier, even though the competence machinery is required to enter the DNA into the cell. In this case, our hypothesis lays on the fact that the DNA taken up by electroporation transiently remains in the periplasmic space, from which it is transported to the cytoplasm by components of the competence apparatus. Therefore, the emission of a putative signal, released as a functional warning during natural competence, may be eluded because in electroporation there is no transport at the external membrane level. This signal would somehow alert of the arrival of unreliable DNA at the cytoplasmic level, eliciting the activation of TtAgo. Actually, our data discard a transcriptional synthesis response as TtAgo's concentration over the growth cycle remained stable and apparently insensitive to the presence of external DNA.

Then, our current hypothesis revolves around the idea of a specific stimulation of TtAgo interference pathway by the entry of DNA by competence. Therefore, TtAgo may rest in an inactive status until a signal is received, probably transmitted from the natural competence apparatus, triggering TtAgo's activation. Whether this activation may happen or not, or if it requires actual contact to the competence apparatus, remains to be elucidated. On the other hand, we hypothesize that TtAgo-mediated interference is defeated during conjugation due to the lack of awareness of the aforementioned "warning call", either because DNA goes directly to the periplasm as it is assumed to happen during electroporation, or if DNA is straightforwardly transported to the cytoplasm. Additional theories on the discriminatory trait such as the protection of the incoming DNA by pilot proteins, its A+T content or the coiling status of the transferred DNA have been initially held up as, according to our current data, TtAgo acts against isogenic genomic DNA and provides similar response on conjugation and electroporation events. However, despite we can determine how the topology of extracellular DNA we ignore its topology during its entrance to the cell.

The ssDNA guides enigma

A final task that it is still pending and worth considering is the unravelling of the generation of the DNA guides, determinant for the DNA interference mediated by TtAgo. According to *in vitro* activity assays, TtAgo itself is able to produce its own ssDNA from a dsDNA template[191]. Conversely, this capacity opposes to our former arguments upon our *in vivo* results. It seems highly unlikely that DNA gets into the cell in a dsDNA conformation, and, on the other hand, transformation assays performed using ssDNA as template provide the same difference between the *ago*$^{+}$ and the *ago*$^{-}$ mutant (Fig. 8.4.C). Hence, we can conclude that *in vivo*, self-generation of the ssDNA guides from dsDNA is not the actual mechanism of guides' generation.

Genome exploration searching for putative candidates for guides' production, pointed out one single protein which could generate them from an ssDNA template. This enzyme, so-called TtPrimPol is encoded in a scarcely conserved chromosomal region of HB27 strain. This polymerase is able to polymerize on DNA *de novo*, with no need for priming. TtPrimPol seems to play a role also in the DNA repair systems, an increased need at high temperatures[63,94,127]. By these means, assuming that our hypothesis linking TtAgo and PrimPol in the same epigenetic pathway is correct, a mutant in PrimPol should provide a similar response to DNA uptake by natural competence as that observed for the *Δago* derivative. Surprisingly, what we detected was that the strain impaired in PrimPol exhibits even higher transformation efficiencies than the *Δago* strain, suggesting that this protein could be involved in further cell defence strategies. However, the double mutant (*ago*$^{-}$*, primPol*$^{-}$) showed an increase in transformability akin to that shown by the in *Δago* mutant, in agreement with our expectations of both proteins cooperating in the defence response mediated by TtAgo. Therefore, these data support our initial hypothesis that TtPrimPol is involved in the TtAgo-mediated DNA-DNA interference. However, the deleterious effect caused by the combination of mutations in TtAgo and TtPrimPol among *T. thermophilus* derivatives was suspicious, inevitably forcing a cautious analysis of this conclusion.

Conclussions

CONCLUSSIONS

The main conclusions derived from this Thesis are the following:

1. *Thermus thermophilus* horizontally transfers DNA through an unconventional and bidirectional conjugation process, which is even more efficient than natural competence.

2. Genes associated to the pTT27 megaplasmid are preferentially transferred (~10 fold) compared to those related to the chromosome.

3. There is no order in the transfer of the genes associated to the pTT27 megaplasmid, whereas in the chromosome there are significant differences in transfer frequencies between loci, suggesting the presence of preferential OriT-like regions.

4. A two-step "*push-pull*" conjugation model is proposed, where DNA donation is independent from its reception, the latter requiring the competence machinery.

5. CptA and HerA, two putative hexamerical ATP-dependent helicases from HB27 strain have been identified as essential for DNA donation. Both show polar localization, suggesting that conjugation takes place at the cell poles.

6. HerA is probably involved in the generation of the DNA to transfer whereas CptA would work in its pumping out from the donor cell, similar to TraB protein from *Streptomyces* spp.

7. The Argonaute protein hinders the entrance of DNA incoming by natural competence, regardless of the DNA template, even if its isogenic or vesicle-protected. However, this barrier is not elicited when the same DNA enters the cell by conjugation or electroporation.

8. The polymerase PrimPol could be contributing with Argonaute in the surveillance system, supplying the required guides to perform the DNA-DNA interference.

CONCLUSIONES

Las principales conclusiones extraídas de esta Tesis Doctoral son:

1. *Thermus thermophilus* transfiere ADN horizontalmente por un sistema no convencional de conjugación bidireccional, más eficiente incluso que la competencia natural.

2. Los genes asociados al megaplásmido pTT27 se transfieren mejor por conjugación (alrededor de 10 veces) que los asociados al cromosoma.

3. No existe un orden de transferencia de los genes del megaplásmido si bien se intuyen múltiples sitios OriT en el cromosoma.

4. Se propone un modelo de conjugación articulado en dos pasos (empuje y captura) donde la donación de ADN es independiente de su recepción, paso éste en el que se requiere la maquinaria de competencia.

5. CptA y HerA, dos posibles helicasas hexámericas dependientes de ATP en la cepa HB27 han sido identificadas como imprescindibles en el proceso de donación de ADN. Su localización polar sugiere que es a través de éstos dónde se produce la conjugación.

6. La proteína HerA está probablemente implicada en la generación del ADN para la transferencia mientras que CptA actuaría en su expulsión, de forma similar a TraB en *Streptomyces* spp.

7. La proteína Argonauta, interfiere en la entrada por competencia natural de todo tipo de DNA en la célula, incluso cuando éste es isogénico o está protegido en vesículas. Sin embargo, esta barrera no se activa cuando el mismo ADN entra por conjugación o por electroporación.

8. La polimerasa PrimPol podría estar actuando con Argonauta en el mecanismo de vigilancia, proporcionándole las guías necesarias para la interferencia.

References

ONLINE RESOURCES

Dr. Averhoff's website_ Online resource for *Thermus thermophilus* competence machinery (http://www.mikrobiologie-frankfurt.de/molecular-microbiology-bioenergetics/prof-dr-beate-averhoff/3-thermus-dna/)
BacMap_ Online databank of bacterial genomes (wishart.biology.ualberta.ca/BacMap)
BMR genomics_ Online platform for oligonucleotides verification (www.bmr-genomics.it)
CELLO_ Subcellular protein localization prediction tool. *v.* 2.5 (http://cello.life.nctu.edu.tw/)
CLUSTWAL Omega_ Protein multiple sequence alignment (http://www.ebi.ac.uk/Tools/msa/clustalo/)
CONJdb_ Online search for T4SS-like proteins (conjdb.web.pasteur.fr)
Darkhorse_ Database of phylogenetically atypical proteins (darkhorse.ucsd.edu)
DOOR^2_ Database enclosing predicted operons in *T. thermophilus* HB27 (csbl.bmb.uga.edu/DOOR)
ELP_ Online software to obtain a synonymous gene (http://mobyle.pasteur.fr/cgi-bin/portal.py -forms::ELP)
EnzymeX_ Software for restriction analysis (nucleobytes.com)
ExPASy_ Resource portal with multiple services such as aminoacid translation, characterization of proteins, etc.(http://www.expasy.org/genomics)
Gene Skew_online portal (genskew.csb.univie.ac.at)
HGT-DB_ Database including the characterization of predicted horizontally transferred genes (http://genomes.urv.cat/HGT-DB/)
HMMer_ Softwre package for sequence analysis (hmmer.janelia.org)
HMMTOP_ Software focused on the prediction of transmembrane helices and topology of proteins (http://www.enzim.hu/hmmtop/)
I-TASSER_ Software executing an hierarchical method for protein structure and function prediction (zhanglab.ccmb.med.umich.edu/I-TASSER)
ISfinder_ Database enclosing the classification and characterization of ISs (is.biotoul.fr)
Island Viewer_ Online software to detect genomic islands (http://www.pathogenomics.sfu.ca/islandviewer/)
KEGG2_ Database of bacterial genomes (http://www.genome.jp/kegg/kegg2.html)
MUSCLE_ Online DNA sequence multiple alignment (http://www.ebi.ac.uk/Tools/msa/muscle/)
NCBI _ Resource for molecular biology information (www.ncbi.nlm.nih.gov)
OligoCalculator_ Online resource employed for G+C calculator, verification fo olingonuclotides (http://www.sciencelauncher.com/oligocalc.html)
OperonDB_ Online resource were predicted operons in *Thermus* spp are characterised (http://www.operondb.jp)
ORF Finder_ Online tool to find open reading frames within a given sequence (http://www.ncbi.nlm.nih.gov/gorf/gorf.html)
PDB_ Online protein data bank (http://www.rcsb.org/pdb/home/home.do)
pFAM_ Online database of protein families (pfam.xfam.org)
Phyre 2.0_ Online portal for protein modelling, prediction and analysis (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index)
PRABI_ Online portal for multiple sequence alignments and BLASTs (doua.prabi.fr)
PSIPred_ Protein Sequence Analysis online software (http://bioinf.cs.ucl.ac.uk/psipred/)
PSORTb_ Subcellular localization prediction tool. *v.* 3.0.2 (http://www.psort.org/psortb/)
RaptorX_ Protein structure prediction server (www.raptorx.uchicago.edu)
Repeat Masker_ Online program to detect insertion sequences by screening interspersed repeats (http://www.repeatmasker.org/)

Repseek_ Analysis tool to detect approximate repeats targeting transposition hotspots (wwwabi.snv.jussieu.fr/public/RepSeek)
SOSUI_ Online resource for secondary structure protein prediction (http://harrier.nagahama-i-bio.ac.jp/sosui/sosui_submit.html)
STRING_ Predicted functional protein interactions and network *v.* 9.0 (string90.embl.de)
SWISS-Model_ Online server employed for protein structure homology-modelling (http://www.swissmodel.expasy.org/)
TMHMM_ Online server which predicts transmembrane helices in proteins *v.* 2.0 (http://www.cbs.dtu.dk/services/TMHMM/)
TMPred_ Database enabling the prediction of membrane-spanning regions and their orientation (http://www.ch.embnet.org/software/TMPRED_form.html)
TN-Pred_ Online server to detect and predict transposases (http://www.mobilomics.cl/index.html)
TopPred_ Online program to predict membrane protein structure (http://mobyle.pasteur.fr/cgi-bin/portal.py? - forms::toppred)
T-Rex_ Online resource to validate and visualize phylogenetic trees (www.trex.uqam.ca)
Uniprot_ Online database of protein information (http://www.uniprot.org/)
Xmipp_ Digital processing of images for high-resolution protein reconstruciton modelling (www.xmipp.cnb.csic.es)

BIBLIOGRAPHY

1 Aagaard, C., Dalgaard, J. Z. & Garrett, R. A. Intercellular mobility and homing of an archaeal rDNA intron confers a selective advantage over intron-cells of *Sulfolobus acidocaldarius*. *Proceedings of the National Academy of Sciences* **92**, 12285-12289 (1995).

2 Acosta, F., Alvarez, L., de Pedro, M. A. & Berenguer, J. Localized synthesis of the outer envelope from *Thermus thermophilus*. *Extremophiles* **16**, 267-275 (2012).

3 Alarico, S. *et al.* Distribution of genes for synthesis of trehalose and mannosylglycerate in *Thermus* spp and direct correlation of these genes with halotolerance. *Applied and Environmental Microbiology* **71**, 2460-2466 (2005).

4 Alvarez, L., Bricio, C., Jose Gomez, M. & Berenguer, J. Lateral transfer of the denitrification pathway genes among *Thermus thermophilus* strains. *Applied and Environmental Microbiology* **77**, 1352-1358 (2011).

5 Alvarez, L., Bricio, C., Blesa, A., Hidalgo, A. & Berenguer, J. The transferable denitrification capability of *Thermus thermophilus*. *Applied and Environmental Microbiology* (2013).

6 Álvarez Muñoz, L. *Ánalisis de la respiración de nitrito en Thermus thermophilus* Microbiology thesis, Universidad Autónoma de Madrid, (2012).

7 Aly, K. A. & Baron, C. The VirB5 protein localizes to the T-pilus tips in *Agrobacterium tumefaciens*. *Microbiology* **153**, 3766-3775 (2007).

8 Amábile-Cuevas, C. F. & Chicurel, M. E. Horizontal Gene Transfer. *American Scientist* **81**, 332-341 (1993).

9 Anthony, K. G., Klimke, W. A., Manchak, J. & Frost, L. S. Comparison of proteins involved in pilus synthesis and mating pair stabilization from the related plasmids F and R100-1: insights into the mechanism of conjugation. *Journal of Bacteriology* **181**, 5149-5159 (1999).

10 Aranda, J., Zinovjev, K., Roca, M. & Tunon, I. Dynamics and reactivity in *Thermus aquaticus* N6-adenine methyltransferase. *Journal of the American Chemical Society* **136**, 16227-16239 (2014).

11 Artimo, P. *et al.* ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Research* **40**, W597-603, doi:10.1093/nar/gks400 (2012).

12 Ashby, M. K. & Bergquist, P. L. Cloning and sequence of *IS1000*, a putative insertion sequence from*Thermus thermophilus* HB8. *Plasmid* **24**, 1-11 (1990).

13 Atmakuri, K., Cascales, E. & Christie, P. J. Energetic components VirD4, VirB11 and VirB4 mediate early DNA transfer reactions required for bacterial type IV secretion. *Molecular Microbiology* **54**, 1199-1211 (2004).

14 Ausubel, F. M. *et al. Current protocols in molecular biology*. (John Wiley and Sons, 1987).

15 Averhoff, B. & Friedrich, A. Type IV pili-related natural transformation systems: DNA transport in mesophilic and thermophilic bacteria. *Archives of Microbiology* **180**, 385-393 (2003).

16 Averhoff, B. Shuffling genes around in hot environments: the unique DNA transporter of *Thermus thermophilus*. *FEMS Microbiol. Rev.* **33**, 611-626 (2009).

17 Barbé, J., Berenguer, J., Casares, L. & Castán, P. Temperature-dependent hypermutational phenotype in recA mutants of *Thermus thermophilus* HB27. (2003).

18 Blas-Galindo, E., Cava, F., López-Vinas, E., Mendieta, J. & Berenguer, J. Use of a dominant rpsL allele conferring streptomycin dependence for positive and negative selection in *Thermus thermophilus*. *Applied and Environmental Microbiology* **73**, 5138-5145 (2007).

19 Blount, Z. D. & Grogan, D. W. New insertion sequences of *Sulfolobus:* functional properties and implications for genome evolution in hyperthermophilic archaea. *Molecular Microbiology* **55**, 312-325 (2005).

20 Bohmert, K. *et al.* AGO1 defines a novel locus of *Arabidopsis* controlling leaf development. *The EMBO journal* **17**, 170-180 (1998).

21 Brazelton, W. J. & Baross, J. A. Abundant transposases encoded by the metagenome of a hydrothermal chimney biofilm. *The ISME journal* **3**, 1420-1424 (2009).
22 Bricio, C., Alvarez, L., Gomez, M. J. & Berenguer, J. Partial and complete denitrification in *Thermus thermophilus*: lessons from genome drafts. *Biochemical Society Transactions* **39**, 249-253 (2011).
23 Bricio, C. *Producción de óxidos de nitrógeno gaseosos en Thermus thermophilus* Microbiology thesis, Autónoma de Madrid, (2012).
24 Brochet, M. *et al.* Atypical association of DDE transposition with conjugation specifies a new family of mobile elements. *Molecular Microbiology* **71**, 948-959 (2009).
25 Brock, T. D. & Freeze, H. *Thermus aquaticus* gen. n. and sp. n., a nonsporulating extreme thermophile. *Journal of Bacteriology* **98**, 289-297 (1969).
26 Brock, T. D. in *Thermophilic microorganisms and life at high temperatures* 72-91 (Springer, 1978).
27 Brouns, S. J. *et al.* Engineering a selectable marker for hyperthermophiles: Crystal structure of a thermostable Bleomycin binding protein. *Journal of Biological Chemistry* (2005).
28 Brüggemann, H. & Chen, C. Comparative genomics of *Thermus thermophilus*: plasticity of the megaplasmid and its contribution to a thermophilic lifestyle. *Journal of Biotechnology* **124**, 654-661 (2006).
29 Burkhardt, J., Vonck, J. & Averhoff, B. Structure and function of PilQ, a secretin of the DNA transporter from the thermophilic bacterium *Thermus thermophilus* HB27. *Journal of Biological Chemistry* **286**, 9977-9984 (2011).
30 Burroughs, A. M., Ando, Y. & Aravind, L. New perspectives on the diversification of the RNA interference system: insights from comparative genomics and small RNA sequencing. *Wiley Interdisciplinary Reviews: RNA* **5**, 141-181 (2014).
31 Bushman, F. Lateral gene transfer: mechanisms and consequences. (Cold Spring Harbor Press, Cold Spring Harbor, 2002).
32 Cabezón, E., Ripoll-Rozada, J., Peña, A., de la Cruz, F. & Arechaga, I. Towards an integrated model of bacterial conjugation. *FEMS Microbiology Reviews* (2014).
33 Cambray, G. & Mazel, D. Synonymous genes explore different evolutionary landscapes. *PLoS Genetics* **4**, e1000256 (2008).
34 Canchaya, C., Fournous, G., Chibani-Chennoufi, S., Dillmann, M.-L. & Brüssow, H. Phage as agents of lateral gene transfer. *Current Opinion in Microbiology* **6**, 417-424 (2003).
35 Canchaya, C., Fournous, G. & Brussow, H. The impact of prophages on bacterial chromosomes. *Molecular Microbiology* **53**, 9 - 18 (2004).
36 Carr, J., Gregory, S. & Dahlberg, A. Transposon mutagenesis of the extremely thermophilic bacterium *Thermus thermophilus* HB27. *Extremophiles*, 1-8 (2014).
37 Carr, J. F., Danziger, M. E., Huang, A. L., Dahlberg, A. E. & Gregory, S. T. Engineering the genome of *Thermus thermophilus* using a counter-selectable marker. *Journal of Bacteriology* (2015).
38 Cascales, E. & Christie, P. J. The versatile bacterial type IV secretion systems. *Nature Reviews Microbiology* **1**, 137-149 (2003).
39 Castan, P. *et al.* The periplasmic space in *Thermus thermophilus*: evidence from a regulation-defective S-layer mutant overexpressing an alkaline phosphatase. *Extremophiles* **6**, 225-232 (2002).
40 Caston, J. R., Berenguer, J., de Pedro, M. A. & Carrascosa, J. L. S-layer protein from *Thermus thermophilus* HB8 assembles into porin-like structures. *Molecular Microbiology* **9**, 65-75 (1993).
41 Cava, F., de Pedro, M. A., Schwarz, H., Henne, A. & Berenguer, J. Binding to pyruvylated compounds as an ancestral mechanism to anchor the outer envelope in primitive bacteria. *Molecular Microbiology* **52**, 677-690 (2004).
42 Cava, F., Zafra, O., Magalon, A., Blasco, F. & Berenguer, J. A new type of NADH dehydrogenase specific for nitrate respiration in the extreme thermophile *Thermus thermophilus*. *Journal of Biological Chemistry* **279**, 45369-45378 (2004).
43 Cava, F. *Control del metabolismo respiratorio en Thermus thermophilus por el elemento genético conjugativo NCE* Biología Molecular thesis, Universidad Autónoma de Madrid, (2007).

44 Cava, F. *et al.* Expression and use of superfolder green fluorescent protein at high temperatures *in vivo*: a tool to study extreme thermophile biology. *Environmental Microbiology* **10**, 605-613 (2008).
45 Cava, F., Hidalgo, A. & Berenguer, J. *Thermus thermophilus* as biological model. *Extremophiles* **13**, 213-231 (2009).
46 Cavalli-Sforza, L. L. Sexuality of bacteria. *Bollettino dell'Istituto sieroterapico milanese* **29**, 281-289 (1950).
47 Cenik, E. S. & Zamore, P. D. Argonaute proteins. *Current Biology* **21**, R446-R449 (2011).
48 César, C. E. *et al.* Unconventional lateral gene transfer in extreme thermophilic bacteria. *International Microbiology* **14**, 187-199 (2012).
49 Chandler, M. *et al.* Breaking and joining single-stranded DNA: the HUH endonuclease superfamily. *Nature Reviews Microbiology* **11**, 525-538 (2013).
50 Chen, I. & Dubnau, D. DNA transport during transformation. *Frontiers in Bioscience* **8**, 544-556 (2003).
51 Chen, I., Christie, P. J. & Dubnau, D. The ins and outs of DNA transfer in bacteria. *Science* **310**, 1456-1460 (2005).
52 Cheng, K. *et al.* Biochemical and functional characterization of the NurA-HerA complex from *Deinococcus radiodurans*. *Journal of Bacteriology* **197**, 2048-2061 (2015).
53 Chimileski, S., Dolas, K., Naor, A., Gophna, U. & Papke, R. T. Extracellular DNA metabolism in *Haloferax volcanii*. *Frontiers in Microbiology* **5** (2014).
54 Chopin, M.-C., Chopin, A. & Bidnenko, E. Phage abortive infection in lactococci: variations on a theme. *Current Opinion in Microbiology* **8**, 473-479 (2005).
55 Christie, P. J. Type IV secretion: intercellular transfer of macromolecules by systems ancestrally related to conjugation machines. *Molecular Microbiology* **40**, 294-305 (2001).
56 Ciccarelli, F. D. *et al.* Toward automatic reconstruction of a highly resolved tree of life. *Science* **311**, 1283-1287 (2006).
57 Claverys, J. P. & Martin, B. Bacterial "competence" genes: signatures of active transformation, or only remnants? *Trends in Microbiology* **11**, 161-165, doi:S0966842X03000647 [pii] (2003).
58 Constantinesco, F., Forterre, P., Koonin, E. V., Aravind, L. & Elie, C. A bipolar DNA helicase gene, herA, clusters with rad50, mre11 and nurA genes in thermophilic archaea. *Nucleic Acids Research* **32**, 1439-1447 (2004).
59 Curtiss, R. I. & Renshaw, J. F+ strains of *Escherichia coli* K-12 defective in *Hfr* formation. *Genetics* **63**, 7 (1969).
60 Da Costa, M. S., Rainey, F. A. & Nobre, F. in *The Prokaryotes* Vol. 7 797-812 (Springer New York, 2006).
61 Da Cunha, V., Guerillot, R., Brochet, M. & Glaser, P. Bacterial integrative mobile genetic elements. (2013).
62 da Silva Rabello, M. C. *et al.* First description of natural and experimental conjugation between Mycobacteria mediated by a linear plasmid. *PloS One* **7** (2012).
63 Daniel, R. M. & Cowan, D. A. Biomolecular stability and life at high temperatures. *Cellular and Molecular Life Sciences* **57**, 250-264 (2000).
64 Danner, D. B., Deich, R. A., Sisco, K. L. & Smith, H. O. An eleven-base-pair sequence determines the specificity of DNA uptake in *Haemophilus* transformation. *Gene* **11**, 311-318 (1980).
65 Davis, B. D. Non-filtrability of the agents of genetic recombination in *Escherichia coli* *Journal of Bacteriology* **60**, 507-508 (1950).
66 de Grado, M., Lasa, I. & Berenguer, J. Characterization of a plasmid replicative origin from an extreme thermophile. *FEMS Microbiology Letters* **165**, 51-57 (1998).
67 De Lorenzo, V., Herrero, M., Jakubzik, U. & Timmis, K. N. Mini-Tn5 transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria. *Journal of Bacteriology* **172**, 6568-6572 (1990).
68 Degryse, E., Glansdorff, N. & Pierard, A. A comparative analysis of extreme thermophilic bacteria belonging to the genus *Thermus*. *Archives of Microbiology* **117**, 189-196 (1978).
69 Demarre, G. *et al.* A new family of mobilizable suicide plasmids based on broad host range R388 plasmid (IncW) and RP4 plasmid (IncPalpha) conjugative machineries and

their cognate *Escherichia coli* host strains. *Research in Microbiology* **156**, 245-255 (2005).
70 Derbyshire, K. M. & Gray, T. A. Distributive Conjugal Transfer: new insights into horizontal gene transfer and genetic exchange in mycobacteria. *Microbiology Spectrum* **2** (2014).
71 Dodsworth, J. A. *et al.* Interdomain conjugal transfer of DNA from bacteria to archaea. *Applied and Environmental Microbiology* **76**, 5644-5647 (2010).
72 Draper, O., Cesar, C. E., Machon, C., de la Cruz, F. & Llosa, M. Site-specific recombinase and integrase activities of a conjugative relaxase in recipient cells. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 16385-16390 (2005).
73 Dubey, G. P. & Ben-Yehuda, S. Intercellular nanotubes mediate bacterial communication. *Cell* **144**, 590-600 (2011).
74 Dubnau, D. DNA uptake in bacteria. *Annual Review of Microbiology* **53**, 217-244 (1999).
75 Elkins, C., Thomas, C. E., Seifert, H. S. & Sparling, P. F. Species-specific uptake of DNA by gonococci is mediated by a 10-base-pair sequence. *Journal of Bacteriology* **173**, 3911-3913 (1991).
76 Ender, C. & Meister, G. Argonaute proteins at a glance. *Journal of Cell Science* **123**, 1819-1823 (2010).
77 Faehnle, C. R., Elkayam, E., Haase, A. D., Hannon, G. J. & Joshua-Tor, L. The making of a slicer: activation of human Argonaute-1. *Cell reports* **3**, 1901-1909 (2013).
78 Finkel, S. E. & Kolter, R. DNA as a nutrient: novel role for bacterial competence gene homologs. *Journal of Bacteriology* **183**, 6288-6293 (2001).
79 Flint, J. L., Kowalski, J. C., Karnati, P. K. & Derbyshire, K. M. The RD1 virulence locus of *Mycobacterium tuberculosis* regulates DNA transfer in *Mycobacterium smegmatis*. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 12598-12603 (2004).
80 Friedrich, A., Hartsch, T. & Averhoff, B. Natural transformation in mesophilic and thermophilic bacteria: identification and characterization of novel, closely related competence genes in *Acinetobacter* sp. strain BD413 and *Thermus thermophilus* HB27. *Applied and Environmental Microbiology* **67**, 3140-3148 (2001).
81 Friedrich, A., Prust, C., Hartsch, T., Henne, A. & Averhoff, B. Molecular analyses of the natural transformation machinery and identification of pilus structures in the extremely thermophilic bacterium *Thermus thermophilus* strain HB27. *Applied and Environmental Microbiology* **68**, 745-755 (2002).
82 Friedrich, A., Rumszauer, J., Henne, A. & Averhoff, B. Pilin-like proteins in the extremely thermophilic bacterium *Thermus thermophilus* HB27: implication in competence for natural transformation and links to type IV pilus biogenesis. *Applied and Environmental Microbiology* **69**, 3695-3700 (2003).
83 Frost, L. S., Leplae, R., Summers, A. O. & Toussaint, A. Mobile genetic elements: the agents of open source evolution. *Nature Reviews Microbiology* **3**, 722-732 (2005).
84 Fürste, J. P., Pansegrau, W., Ziegelin, G., Kröger, M. & Lanka, E. Conjugative transfer of promiscuous IncP plasmids: interaction of plasmid-encoded products with the transfer origin. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 1771-1775 (1989).
85 Furuta, Y., Abe, K. & Kobayashi, I. Genome comparison and context analysis reveals putative mobile forms of restriction–modification systems and related rearrangements. *Nucleic Acids Research*, 1226 (2010).
86 Garcillán-Barcia, M. P. & de la Cruz, F. Why is entry exclusion an essential feature of conjugative plasmids? *Plasmid* **60**, 1-18 (2008).
87 Gogarten, J. P. & Townsend, J. P. Horizontal gene transfer, genome innovation and evolution. *Nature Reviews Microbiology* **3**, 679-687 (2005).
88 Gold, V. A., Salzer, R., Averhoff, B. & Kühlbrandt, W. Structure of a type IV pilus machinery in the open and closed state. *eLife* **4**, e07380, doi:10.7554/eLife.07380 (2015).
89 Goldfarb, T. *et al.* BREX is a novel phage resistance system widespread in microbial genomes. *The EMBO journal* **34**, 169-183 (2015).
90 Gomis-Rüth, F. X. *et al.* The bacterial conjugation protein TrwB resembles ring helicases and F1-ATPase. *Nature* **409**, 637-641 (2001).

91 Gregory, S. T. & Dahlberg, A. E. Transposition of an insertion sequence, *ISTth7,* in the genome of the extreme thermophile *Thermus thermophilus* HB8. *FEMS Microbiology Letters* **289**, 187-192 (2008).
92 Grissa, I., Vergnaud, G. & Pourcel, C. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics* **8**, 172 (2007).
93 Grogan, D. W. Exchange of genetic markers at extremely high temperatures in the archaeon *Sulfolobus acidocaldarius*. *Journal of Bacteriology* **178**, 3207-3211 (1996).
94 Grogan, D. W. The question of DNA repair in hyperthermophilic archaea. *Trends in Microbiology* **8**, 180-185 (2000).
95 Guglielmini, J., Quintais, L., Garcillán-Barcia, M. P., De La Cruz, F. & Rocha, E. P. The repertoire of ICE in prokaryotes underscores the unity, diversity, and ubiquity of conjugation. *PLoS Genetics* **7**, e1002222 (2011).
96 Gupta, R. S. The natural evolutionary relationships among prokaryotes. *Critical Reviews in Microbiology* **26**, 111-131 (2000).
97 Hanahan, D. Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology* **166**, 557-580 (1983).
98 Henne, A. *et al.* The genome sequence of the extreme thermophile *Thermus thermophilus*. *Nature Biotechnology* **22**, 547 - 553 (2004).
99 Herrero, M., de Lorenzo, V. & Timmis, K. N. Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in gram-negative bacteria. *Journal of Bacteriology* **172**, 6557-6567 (1990).
100 Höck, J. & Meister, G. The Argonaute protein family. *Genome Biology* **9**, 210 (2008).
101 Horvath, P. & Barrangou, R. CRISPR/Cas, the immune system of bacteria and archaea. *Science* **327**, 167-170 (2010).
102 Hoshino, T. *et al.* Isolation and partial characterization of carotenoid underproducing and overproducing mutants from an extremely thermophilic *Thermus thermophilus* HB27. *Journal of Fermentation and Bioengineering* **77**, 131-136 (1994).
103 Hur, J. K., Olovnikov, I. & Aravin, A. A. Prokaryotic Argonautes defend genomes against invasive DNA. *Trends in Biochemical Sciences* **39**, 257-259 (2014).
104 Inoue, H., Nojima, H. & Okayama, H. High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96**, 23-28 (1990).
105 Jain, R., Rivera, M. C., Moore, J. E. & Lake, J. A. Horizontal gene transfer accelerates genome innovation and evolution. *Molecular Biology and Evolution* **20**, 1598-1602 (2003).
106 Jinek, M. & Doudna, J. A. A three-dimensional view of the molecular machinery of RNA interference. *Nature* **457**, 405-412 (2009).
107 Johnston, C., Martin, B., Fichant, G., Polard, P. & Claverys, J.-P. Bacterial transformation: distribution, shared mechanisms and divergent control. *Nature Reviews Microbiology* **12**, 181-196 (2014).
108 Kadurugamuwa, J. L. & Beveridge, T. J. Bacteriolytic effect of membrane vesicles from *Pseudomonas aeruginosa* on other bacteria including pathogens: conceptually new antibiotics. *Journal of Bacteriology* **178**, 2767-2774 (1996).
109 Klostermeier, D. Rearranging RNA structures at 75° C? toward the molecular mechanism and physiological function of the thermus thermophilus DEAD‐box helicase hera. *Biopolymers* **99**, 1137-1146 (2013).
110 Koerdt, A., Gödeke, J., Berger, J., Thormann, K. M. & Albers, S.-V. Crenarchaeal biofilm formation under extreme conditions. *PLoS ONE* **5**, e14104 (2010).
111 Koonin, E., Makarova, K. & L., A. Horizontal gene transfer in prokaryotes - quantification and classification. *Annual Review of Microbiology* (2001).
112 Kosono, S., Kataoka, M., Seki, T. & Yoshida, T. The TraB protein, which mediates the intermycelial transfer of the *Streptomyces* plasmid pSN22, has functional NTP‐binding motifs and is localized to the cytoplasmic membrane. *Molecular Microbiology* **19**, 397-405 (1996).
113 Koyama, Y., Hoshino, T., Tomizuka, N. & Furukawa, K. Genetic transformation of the extreme thermophile *Thermus thermophilus* and of other *Thermus* spp. *Journal of Bacteriology* **166**, 338-340 (1986).

114 Kumwenda, B., Litthauer, D. & Reva, O. Analysis of genomic rearrangements, horizontal gene transfer and role of plasmids in the evolution of industrial important *Thermus* species. *BMC Genomics* **15**, 813 (2014).
115 Labrie, S. J., Samson, J. E. & Moineau, S. Bacteriophage resistance mechanisms. *Nature Reviews Microbiology* **8**, 317-327 (2010).
116 Lang, A. S. & Beatty, J. T. Importance of widespread gene transfer agent genes in α-Proteobacteria. *Trends in Microbiology* **15**, 54-62 (2007).
117 Lang, A. S., Zhaxybayeva, O. & Beatty, J. T. Gene transfer agents: phage-like elements of genetic exchange. *Nature Reviews Microbiology* **10**, 472-482 (2012).
118 Lasa, I., Caston, J. R., Fernandez-Herrero, L. A., de Pedro, M. A. & Berenguer, J. Insertional mutagenesis in the extreme thermophilic eubacteria *Thermus thermophilus* HB8. *Molecular Microbiology* **6**, 1555-1564 (1992).
119 Lawrence, J. G. Gene transfer, speciation, and the evolution of bacterial genomes. *Current Opinion in Microbiology* **2**, 519-523 (1999).
120 Lennox, E. Transduction of linked genetic characters of the host by bacteriophage P1. *Virology* **1**, 190-206 (1955).
121 Lerat, E., Daubin, V., Ochman, H. & Moran, N. A. Evolutionary origins of genomic repertoires in bacteria. *PLoS Biology* **3**, e130 (2005).
122 Llosa, M., Gomis-Rüth, F. X., Coll, M. & Cruz, F. d. l. Bacterial conjugation: a two-step mechanism for DNA transport. *Molecular Microbiology* **45**, 1-8 (2002).
123 Lossouarn, J. *et al.* An abyssal mobilome: viruses, plasmids and vesicles from deep-sea hydrothermal vents. *Research in Microbiology* (2015).
124 Ludtke, S. J., Baldwin, P. R. & Chiu, W. EMAN: semiautomated software for high-resolution single-particle reconstructions. *Journal of Structural Biology* **128**, 82-97 (1999).
125 Luo, Y. & Wasserfallen, A. Gene transfer systems and their applications in Archaea. *Systematic and Applied Microbiology* **24**, 15-25 (2001).
126 Mahillon, J. & Chandler, M. Insertion Sequences. *Microbiology and Molecular Biology Reviews* **62**, 725-774 (1998).
127 Makarova, K., Aravind, L., Grishin, N., Rogozin, I. & Koonin, E. A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. *Nucleic Acids Research* **30**, 482 - 496 (2002).
128 Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I. & Koonin, E. V. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biology Direct* **1**, 7 (2006).
129 Makarova, K. S., Wolf, Y. I., Van der Oost, J. & Koonin, E. V. Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements. *Biology Direct* **4**, 29 (2009).
130 Makarova, K. S., Wolf, Y. I. & Koonin, E. V. Comparative genomics of defense systems in archaea and bacteria. *Nucleic Acids Research* **41**, 4360-4377 (2013).
131 Manaia, C. M. *et al.* Halotolerant *Thermus* strains from marine and terrestrial hot-springs belong to *Thermus thermophilus* (ex Oshima and Imahori, 1974) nom. rev. emend. *Systematic and Applied Microbiology* **17**, 526-532 (1995).
132 Marguet, E. *et al.* Membrane vesicles, nanopods and/or nanotubes produced by hyperthermophilic archaea of the genus *Thermococcus*. *Biochemical Society Transactions* **41** (2013).
133 Marmur, J. A procedure for the isolation of deoxyribonucleic acid from micro-organisms. *Journal of Molecular Biology* **3**, 208-IN201 (1961).
134 McDaniel, L. D. *et al.* High frequency of horizontal gene transfer in the oceans. *Science* **330**, 50-50 (2010).
135 McDaniel, L. D., Young, E. C., Ritchie, K. B. & Paul, J. H. Environmental factors influencing Gene Transfer Agent (GTA) mediated transduction in the subtropical ocean. *PLoS ONE* **7**, e43506 (2012).
136 Mell, J. C., Hall, I. M. & Redfield, R. J. Defining the DNA uptake specificity of naturally competent *Haemophilus influenzae* cells. *Nucleic Acids Research* **40**, 8536-8549 (2012).
137 Miller, D. N., Bryant, J. E., Madsen, E. L. & Ghiorse, W. C. Evaluation and optimization of DNA extraction and urification procedures for soil and sediment samples. *Applied and Environmental Microbiology* **65**, 4715-4724 (1999).

138 Möller, C. & Van Heerden, E. Isolation of a soluble and membrane-associated Fe (III) reductase from the thermophile, *Thermus scotoductus* (SA-01). *FEMS Microbiology Letters* **265**, 237-243 (2006).
139 Moreno, R., Zafra, O., Cava, F. & Berenguer, J. Development of a gene expression vector for *Thermus thermophilus* based on the promoter of the respiratory nitrate reductase. *Plasmid* **49**, 2-8 (2003).
140 Mortimer, T. D. & Pepperell, C. S. Genomic signatures of distributive conjugal transfer among Mycobacteria. *Genome Biology and Evolution* **6**, 2489-2500 (2014).
141 Nakamura, A., Takakura, Y., Kobayashi, H. & Hoshino, T. *In vivo* directed evolution for thermostabilization of *Escherichia coli* hygromycin B phosphotransferase and the use of the gene as a selection marker in the host-vector system of *Thermus thermophilus*. *Journal of Bioscience and Bioengineering* **100**, 158-163 (2005).
142 Naor, A. & Gophna, U. Cell fusion and hybrids in Archaea: Prospects for genome shuffling and accelerated strain development for biotechnology. *Bioengineered* **4**, 126-129 (2013).
143 Nelson, K. *et al.* Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. *Nature* **399**, 323 - 329 (1999).
144 Nielsen, K. M., Ray, J. L. & Johnsen, P. J. in *Encyclopedia of Microbiology (Third Edition)* (ed Schaechter Editor-in-Chief: Moselio) 587-596 (Academic Press, 2009).
145 O'Toole, G. A. Microtiter Dish Biofilm Formation Assay. *Journal of Visualized Experiments : JoVE*, 2437, doi:10.3791/2437 (2011).
146 Ochman, H., Lawrence, J. G. & Groisman, E. A. Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**, 299-304 (2000).
147 Ochman, H. & Davalos, L. M. The nature and dynamics of bacterial genomes. *Science* **311**, 1730-1733 (2006).
148 Olovnikov, I., Chan, K., Sachidanandam, R., Newman, D. K. & Aravin, A. A. Bacterial argonaute samples the transcriptome to identify foreign DNA. *Molecular Cell* **51**, 594-605 (2013).
149 Omelchenko, M., Makarova, K., Wolf, Y., Rogozin, I. & Koonin, E. Evolution of mosaic operons by horizontal gene transfer and gene displacement *in situ*. *Genome Biology* **4**, R55 (2003).
150 Omelchenko, M. *et al.* Comparative genomics of *Thermus thermophilus* and *Deinococcus radiodurans*: divergent routes of adaptation to thermophily and radiation resistance. *BMC Evolutionary Biology* **5**, 57 (2005).
151 Oren, A. DNA as genetic material and as a nutrient in halophilic Archaea. Commentary on “Extracellular DNA metabolism in Haloferax volcanii” by S. Chimileski et al. (Frontiers in Microbiology 5, 57). *Frontiers in Microbiology* **5**, doi:10.3389/fmicb.2014.00539 (2014).
152 Oshima, T. Unique polyamines produced by an extreme thermophile, *Thermus thermophilus*. *Amino Acids* **33**, 367 - 372 (2007).
153 Palchevskiy, V. & Finkel, S. E. *Escherichia coli* competence gene homologs are essential for competitive fitness and the use of DNA as a nutrient. *Journal of Bacteriology* **188**, 3902-3910, doi:10.1128/jb.01974-05 (2006).
154 Peng, Y., Cai, J., Wang, W. & Su, B. Multiple inter-kingdom horizontal gene transfers in the evolution of the phosphoenolpyruvate carboxylase gene family. *PLoS ONE* **7**, e51159 (2012).
155 Pérez-Mendoza, D. & de la Cruz, F. Escherichia coli genes affecting recipient ability in plasmid conjugation: Are there any? *BMC Genomics* **10**, 71 (2009).
156 Prangishvili, D. *et al.* Conjugation in archaea: frequent occurrence of conjugative plasmids in *Sulfolobus*. *Plasmid* **40**, 190-202 (1998).
157 Ramirez-Arcos, S., Fernandez-Herrero, L. A., Marin, I. & Berenguer, J. Anaerobic growth, a property horizontally transferred by an *Hfr*-like mechanism among extreme thermophiles. *Journal of Bacteriology* **180**, 3137-3143 (1998).
158 Redfield, R. J. Genes for breakfast: the have-your-cake-and-eat-it-too of bacterial transformation. *Journal of Heredity* **84**, 400-404 (1993).
159 Renelli, M., Matias, V., Lo, R. Y. & Beveridge, T. J. DNA-containing membrane vesicles of *Pseudomonas aeruginosa* PAO1 and their genetic transformation potential. *Microbiology* **150**, 2161-2169 (2004).
160 Reuther, J., Gekeler, C., Tiffert, Y., Wohlleben, W. & Muth, G. Unique conjugation mechanism in mycelial streptomycetes: a DNA-binding ATPase translocates

unprocessed plasmid DNA at the hyphal tip. *Molecular Microbiology* **61**, 436-446 (2006).
161 Rivera, N. *Termoestabilización de proteínas de interés biotecnológico* Microbiología thesis, Universidad Autónoma de Madrid, (2013).
162 Rosenberg, A. H. *et al.* Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene* **56**, 125-135 (1987).
163 Rosenshine, I., Tchelet, R. & Mevarech, M. The mechanism of DNA transfer in the mating system of an archaebacterium. *Science* **245**, 1387-1389 (1989).
164 Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *Journal of Molecular Biology* **333**, 721-745 (2003).
165 Rumszauer, J., Schwarzenlander, C. & Averhoff, B. Identification, subcellular localization, and functional interactions of PilMNOWQ and PilA4 involved in transformation competency and pilus biogenesis in the thermophilic bacterium *Thermus thermophilus* HB27. *FEBS Journal* **273**, 3261-3272 (2006).
166 Sallam, K. I., Mitani, Y. & Tamura, T. Construction of random transposition mutagenesis system in *Rhodococcus erythropolis* using *IS1415*. *Journal of Biotechnology* **121**, 13-22 (2006).
167 Salzer, R. *et al.* The DNA uptake ATPase PilF of *Thermus thermophilus*: a reexamination of the zinc content. *Extremophiles*, 1-2 (2013).
168 Salzer, R., Joos, F. & Averhoff, B. Type IV pilus biogenesis, twitching motility and DNA uptake in *Thermus thermophilus*: Discrete roles of antagonistic ATPases PilF, PilT1 and PilT2. *Applied and Environmental Microbiology* (2013).
169 Salzer, R., Kern, T., Joos, F. & Averhoff, B. Environmental factors affecting the expression of type IV pilus genes as well as piliation of *Thermus thermophilus*. *FEMS Microbiology Letters* **357**, 56-62 (2014).
170 Sambrook, J. & Russell, D. W. *Molecular Cloning: A Laboratory Manual*. (Cold Spring Harbor Laboratory Press, 2001).
171 Sasaki, H. M. & Tomari, Y. The true core of RNA silencing revealed. *Nature Structural & Molecular Biology* **19**, 657-660 (2012).
172 Schell, J. & Van Montagu, M. Transfer, maintenance, and expression of bacterial Ti-plasmid DNA in plant cells transformed with *A. tumefaciens*. *Brookhaven Symposia in Biology*, 36-49 (1977).
173 Schiestl, R. H. & Petes, T. D. Integration of DNA fragments by illegitimate recombination in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 7585-7589 (1991).
174 Schwarzenlander, C. & Averhoff, B. Characterization of DNA transport in the thermophilic bacterium *Thermus thermophilus* HB27. *FEBS Journal* **273**, 4210-4218 (2006).
175 Seed, K. D. Battling phages: how bacteria defend against viral attack. *PLoS pathogens* **11**, e1004847 (2015).
176 Sen, S. & Oriel, P. *Transfer of transposon Tn916 from Bacillus subtilis to Thermus aquaticus*. Vol. 67 (1990).
177 Shabalina, S. A. & Koonin, E. V. Origins and evolution of eukaryotic RNA interference. *Trends in Ecology & Evolution* **23**, 578-587 (2008).
178 Sheng, G. *et al.* Structure-based cleavage mechanism of *Thermus thermophilus* Argonaute DNA guide strand-mediated DNA target cleavage. *Proceedings of the National Academy of Sciences of the United States of America* (2013).
179 Shetty, A., Chen, S., Tocheva, E. I., Jensen, G. J. & Hickey, W. J. Nanopods: a new bacterial structure and mechanism for deployment of outer membrane vesicles. *PLoS ONE* **6**, e20725 (2011).
180 Shetty, A. & Hickey, W. J. Effects of outer membrane vesicle formation, surface-layer production and nanopod development on the metabolism of phenanthrene by *Delftia acidovorans* Cs1-4. *PLoS ONE* **9**, e92143 (2014).
181 Smillie, C., Garcillan-Barcia, M. P., Francia, M. V., Rocha, E. P. & de la Cruz, F. Mobility of plasmids. *Microbiology and Molecular Biology Reviews* **74**, 434-452 (2010).
182 Smith, H. O., Gwinn, M. L. & Salzberg, S. L. DNA uptake signal sequences in naturally transformable bacteria. *Research in Microbiology* **150**, 603-616 (1999).

183 Soler, N., Marguet, E., Verbavatz, J.-M. & Forterre, P. Virus-like vesicles and extracellular DNA produced by hyperthermophilic archaea of the order *Thermococcales*. *Research in Microbiology* **159**, 390-399 (2008).
184 Solomon, J. M. & Grossman, A. D. Who's competent and when: regulation of natural genetic competence in bacteria. *Trends in Genetics* **12**, 150-155 (1996).
185 Song, J.-J., Smith, S. K., Hannon, G. J. & Joshua-Tor, L. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* **305**, 1434-1437 (2004).
186 Sowers, K. R. & Schreier, H. J. Gene transfer systems for the Archaea. *Trends in Microbiology* **7**, 212-219 (1999).
187 Staals, Raymond H. J. *et al.* RNA Targeting by the Type III-A CRISPR-Cas Csm Complex of *Thermus thermophilus*. *Molecular Cell* (2014).
188 Stedman, K. M. *et al.* pING family of conjugative plasmids from the extremely thermophilic archaeon *Sulfolobus islandicus*: insights into recombination and conjugation in Crenarchaeota. *Journal of Bacteriology* **182**, 7014-7020 (2000).
189 Swarts, D. C. *et al.* DNA-guided DNA interference by a prokaryotic Argonaute. *Nature* **507**, 258-261 (2014).
190 Swarts, D. C. *et al.* The evolutionary journey of Argonaute proteins. *Nature Structural & Molecular Biology* **21**, 743-753 (2014).
191 Swarts, D. C. *Evolution, role and mechanism of prokaryotic Argonaute porteins* PhD thesis, Wageningen university, NL, (2015).
192 Tanaka, T., Kawano, N. & Oshima, T. Cloning of 3-isopropylmalate dehydrogenase gene of an extreme thermophile and partial purification of the gene product. *Journal of Biochemistry* **89**, 677-682 (1981).
193 Tatum, E. L. & Lederberg, J. Gene recombination in the bacterium *Escherichia coli*. *Journal of Bacteriology* **53**, 673-684 (1947).
194 Terns, M. P. & Terns, R. M. CRISPR-based adaptive immune systems. *Current Opinion in Microbiology* **14**, 321-327 (2011).
195 Terns, R. M. & Terns, M. P. CRISPR-based technologies: prokaryotic defense weapons repurposed. *Trends in Genetics* **30**, 111-118 (2014).
196 Thoma, L. & Muth, G. Conjugative DNA transfer in *Streptomyces* by TraB: is one protein enough? *FEMS Microbiology Letters* **337**, 81-88 (2012).
197 Thoma, L. & Muth, G. The conjugative DNA-transfer apparatus of *Streptomyces*. *International Journal of Medical Microbiology* **305**, 224-229 (2015).
198 Thomas, C. M. & Nielsen, K. M. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nature Reviews Microbiology* **3**, 711-721 (2005).
199 Tock, M. R. & Dryden, D. T. The biology of restriction and anti-restriction. *Current Opinion in Microbiology* **8**, 466-472 (2005).
200 Toussaint, A. & Chandler, M. in *Bacterial Molecular Networks* 57-80 (Springer, 2012).
201 Ummels, R. *et al.* Identification of a novel conjugative plasmid in mycobacteria that requires both type IV and type VII secretion. *mBio* **5**, e01744-01714 (2014).
202 Utsumi, R. *et al.* Isolation and characterization of the IS3-like element from *Thermus aquaticus*. *Bioscience, biotechnology, and biochemistry* **59**, 1707-1711 (1995).
203 van der Oost, J., Westra, E. R., Jackson, R. N. & Wiedenheft, B. Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nature Reviews Microbiology* **12**, 479-492 (2014).
204 Vieira, J. & Messing, J. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**, 259-268 (1982).
205 Vogelmann, J. *et al.* Conjugal plasmid transfer in *Streptomyces* resembles bacterial chromosome segregation by FtsK/SpoIIIE. *The EMBO journal* **30**, 2246-2254 (2011).
206 Wang, J., Parsons, L. M. & Derbyshire, K. M. Unconventional conjugal DNA transfer in Mycobacteria. *Nature Genetics* **34**, 80-84 (2003).
207 Wang, J., Karnati, P. K., Takacs, C. M., Kowalski, J. C. & Derbyshire, K. M. Chromosomal DNA transfer in *Mycobacterium smegmatis* is mechanistically different from classical *Hfr* chromosomal DNA transfer. *Molecular Microbiology* **58**, 280-288 (2005).
208 Wang, Y., Sheng, G., Juranek, S., Tuschl, T. & Patel, D. J. Structure of the guide-strand-containing argonaute silencing complex. *Nature* **456**, 209-213 (2008).
209 Wei, K. *et al.* Argonaute protein as a linker to command center of physiological processes. *Chinese Journal of Cancer Research* **25**, 430-441 (2013).

210 Weisburg, W., Giovannoni, S. & Woese, C. The *Deinococcus-Thermus* phylum and the effect of rRNA composition on phylogenetic tree construction. *Systematic and Applied Microbiology* **11**, 128 - 134 (1989).
211 Westra, E. R. *Adaptative and heritable inmunity in bacteria* PhD thesis, Wageningen University,NL, (2013).
212 Willkomm, S., Zander, A., Gust, A. & Grohmann, D. A prokaryotic twist on argonaute function. *Life* **5**, 538-553 (2015).
213 Wozniak, R. A. & Waldor, M. K. Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nature Reviews Microbiology* **8**, 552-563 (2010).
214 Yu, M. X., Slater, M. R. & Ackermann, H. W. Isolation and characterization of *Thermus* bacteriophages. *Archives of Virology* **151**, 663-679 (2006).
215 Zafra, O. *et al.* A cytochrome *c* encoded by the *nar* operon is required for the synthesis of active respiratory nitrate reductase in *Thermus thermophilus*. *FEBS Letters* **523**, 99-102 (2002).
216 Zechner, E. *et al.* in *The horizontal gene pool: bacterial plasmids and gene spread* (ed Christopher M. Thomas) 87-174 (Hardwood Academic, 2000).
217 Zhu, Z. *et al.* Characterization of cleavage intermediate and star sites of RM. *Tth111II*. *Scientific Reports* **4** (2014).
218 Zylicz-Stachula, A. *et al.* Related bifunctional restriction endonuclease-methyltransferase triplets: *TspDTI, Tth111II/TthHB27I* and *TsoI* with distinct specificities. *BMC Molecular Biology* **13**, 13 (2012).

Annexes

ANNEX I

TABLE I. Plasmids employed and constructed in this dissertation

Plasmid	Description/Use	Reference
pET28b(+)	Kan^R, lacI, *E. coli* expression gene dependent on Ö10 promotor from RNA polymerase of T7 phague. Includes fusion of a His-tag (6 Histidines) at N-terminal site. Protein overexpression for purification	Novagen
pUC19/18	Amp^R, P-lac-lacZ'. Cloning	Vieira and Messing[204]
pUC119	Amp^R, P-lac-lacZ'. Cloning	Vieira and Messing[204]
pUC18*NotI*	Amp^R, P-lac-lacZ'. Includes a *NotI* site for cloning	Berenguer's laboratory
pUC19*::kat*	Amp^R, Kan^R. pUC19 derivative with Kan resistance cassette cloned at *XbaI* site. Cloning	This work
pUC19*::hyg*	Amp^R, Hyg^R. pUC19 derivative with Hyg resistance cassette inserted at *XbaI* site. Cloning	This work
pK18 (pK118)	Kan^R. Suicide vector for mutagenesis. Cloning, generation of single insertion mutants	Cava *et al*[42]
pH118	Hyg^R. Suicide vector for mutagenesis. Cloning, generation of single insertion mutants	Álvarez[6]
pMK184	Kan^R. Cloning in *T. thermophilus*	Cava *et al*[43]
pMH184	Hyg^R. Cloning in *T. thermophilus*	Cava *et al*[43]
pMH185	Hyg^R. Cloning in *T. thermophilus*. pMH184 derivative with no *BamHI* nor *NdeI* sites in the *hph5* cassette	Dr. Hidalgo, unpublished
pMH118	Hyg^R. Cloning in *T. thermophilus*	Berenguer's laboratory
pWUR	$Bleo^R$. Cloning. Transformation and expression in *T. thermophilus*	Brouns *et al*[27]
pZEROBlunt TOPO	Amp^R. Cloning. Long-term storage of DNA fragments	Thermo Fisher
pK118 Δ*nirS*	Kan^R.Derivative from pK118 harbouring *nirS* gene, NirS mutant (XbaI/EcoRI)	Álvarez[6]
pUC19::*TTC 0313::hyg*	Hyg^R, Amp^R. pUC19 derivative harbouring *TTC0313* gene (XbaI/EcoRI), interrupted by *hph5* resistance cassette	Álvarez[6]
pnarC::kat	Kan^R. *narC* gene interrupted by the insertion of *kat* resistance cassette	Zafra *et al*[215]
pUC19Δ*norB ::kat*	Amp^R,Kan^R. pUC19 derivative harbouring upstream and downstream flanking regions of *norB* gene. *norB* deletion	Bricio[23]
pMHPnqosY FP	Hyg^R. pMH184 derivative harbouring the reporter superfolder YFP (sYFP), controlled by the promoter Pnqo (XbaI/HindIII).*T. thermophilus* replicative, Protein Fusion mutants	Cava *et al*[44]
pMKPnqosY FP	Kan^R. *T. thermophilus* replicative. pMKnqobgaA derivative where the *bgaA* reporter has been replaced for the sYFP	Cava[43]
pMK184/Taq	Kan^R. pMK184 plasmid extracted from *Thermus aquaticus. Dam-*	This work
p#36	**porf333::kat**. Kan^R. pBIISK derivative harbouring *TTC1620* (*pilT*) disrupted by *kat* cassette (EcoRI/XbaI)	Salzer *et al*[169]
p#186	**pUC19::Δ*pilF::*Kat**. Amp^R, Kan^R. *pilF* gene (*TTC1622*) is replaced by upstream and downstream flanking sequences	Salzer *et al*[169]
p#201	**pBSKpilOWbleoChor**. $Bleo^R$. Deletion of *pilQ* (*TTC1017*)	Burkhardt *et al*[29]
p#27	**pAY66**. Amp^R, Kan^R. pTZ19R derivative harbouring *orf1388* (*pilT*) interrupted by the *kat* cassette	Friedrich *et al*[82]
p#255	**pUC19::ΔpilF+T::Bleo**. Amp^R, $Bleo^R$. Deletion of *pilF* and *pilT (TTC1621-1622)*, replaced by upstream and downstream sequences of the deleted region	Salzer *et al*[168]
p#185	**pUC19::ΔpilF::Bleo**. Amp^R, $Bleo^R$. Deletion of *pilF* gene (*TTC1622*), replaced by *bleo* cassette (EcoRI/HindIII)	Salzer *et al*[168]

pHrecA::kat	**pH118::Δ*recA::kat*.** Hyg^R, Kan^R. Deletion of *recA* gene, replaced by *kat*, flanked by upstream and downstream sequences of the *recA* gene	Rivera[161]
pUC18XX:H	**pUC19:: molib-ribo::*hyg.*** Amp^R, Hyg^R. pUC19 derivative harbouring upstream region of *TTC1836* (molib-tungst) and downstream of *TTC1847* (ribo), interpersed by *hyg* cassette. Deletion of prepilin cluster *TTC1836-1847*	This work
pNCY::*rpmB* ::KAT	Amp^R, Kan^R. pNCY derivative encoding the *kat* cassette which expression is controlled by the promoter and terminator of *T. thermophilus rpmB* gene	Dr. Hidalgo, unpublished
pΔ42::kat	**pUC119::ΔTTP042::*kat*.** Amp^R, Kan^R. pUC19 derivative harbouring upstream and dowstream sequences of *TTP042*, which has been replaced by *kat*	Berenguer's laboratory
pSW23T	**pSW23::oriTRP4.** Cm^R. oriVR6Kγ. Transposase mutagenesis background cloning vector	Demarre *et al*[69]
pET11a::*TTHA0522*	Amp^R. Expression vector encoding *T. thermophilus* HB8 gene *TTHA0522*, tailed with a His-tag at N-terminal (BglII/EcoRI)	RIKEN BioRseource Center
pCEC7	**pUC18:: ΔpAGO::kat.** Kan^R. pUC18 derivative harbouring upstream and dowsntream regions of *ttAgo* gene (*TTP026*), interpersed by the *kat* cassette (HindIII/EcoRI)	C. Elvira, unpublished
pCEC13	**pH118::AGOdown/up**. Hyg^R. pH118 derivative harbouring upstream and dowsntream regions of *ttAgo* gene (*TTP026*) (HindIII/EcoRI)	C. Elvira, unpublished
pCEC30	**pK18::TTP167**. Kan^R. Insertion of the *kat* cassette interrupting *TTP167* gene (EcoRI/HindIII)	C. Elvira. This work
pCEC31	**pK18::TTP219**. Kan^R. Insertion of the *kat* cassette interrupting *TTP219* gene (EcoRI/HindIII)	C. Elvira. This work
pCEC32	**pH118::TTP167**. Hyg^R. Insertion of the *hph5* cassette interrupting *TTP167* gene (EcoRI/HindIII)	C. Elvira. This work
pCEC33	**pK18::TTC1415.** Kan^R. Insertion of the *kat* cassette interrupting *TTC1415* gene (EcoRI/HindIII)	C. Elvira. This work
pCEC34	**pK18::TTC1844.** Kan^R. Insertion of the *kat* cassette interrupting *TTC1844* gene (EcoRI/HindIII)	C. Elvira. This work
pCEC35	**pK18::TTC0638**. Kan^R. Insertion of the *kat* cassette interrupting *TTC0638* gene (EcoRI/HindIII)	C. Elvira. This work
pCEC36	**pK18::TTC0893**. Kan^R. Insertion of the *kat* cassette interrupting TTC0893 gene (EcoRI/HindIII)	C. Elvira. This work
pCEC37	**pK18::TTC1430**. Kan^R. Insertion of the kat cassette interrupting *TTC1430* gene (EcoRI/HindIII)	C. Elvira. This work
pCEC40	**pH118::pilQ.** Hyg^R. Insertion of the *hph5* cassette interrupting *TTC1017* gene (*pilQ*) (EcoRI/HindIII)	C. Elvira. This work
pCEC42	**pH118::TTC1844.** Hyg^R. Insertion of the *hph5* cassette interrupting *TTC1844* gene (EcoRI/HindIII)	C. Elvira. This work
pCBG1	**pH118::TTP046**. Hyg^R. Insertion of the *hph5* cassette interrupting *TTP046* gene (PsiI/XmnI)	Bricio[23]
pCBG2	**pK18::TTP046**. Kan^R. Insertion of the *kat* cassette interrupting *TTP046* gene (PsiI/XmnI)	Bricio[23]
pCBG3	**pUC119::ΔTTP146::hyg**. Amp^R, Hyg^R. pUC19 derivative harbouring upstream and dowstream sequences of *TTP146*, which has been replaced by *hyg* cassettte (BglII)	Bricio[23]
pAB3	**pK18::TTP084**. Kan^R. Insertion of the *kat* cassette interrupting *TTP084* gene (EcoRI/HindIII)	This work
pAB4	**pK18::TTP128**. Kan^R. Insertion of the *kat* cassette interrupting *TTP128* gene (EcoRI/HindIII)	This work
pAB5	**pK18::TTP208**. Kan^R. Insertion of the *kat* cassette interrupting *TTP208* gene (EcoRI/HindIII)	This work
pAB6	**pK18::TTP191**. Kan^R. Insertion of the *kat* cassette interrupting *TTP191* gene (EcoRI/HindIII)	This work
pAB7	**pK18::TTP081**. Kan^R. Insertion of the *kat* cassette interrupting *TTP081* gene (EcoRI/HindIII)	This work
pAB8	**pK18::TTP085**. Kan^R. Insertion of the *kat* cassette interrupting *TTP085* gene (EcoRI/HindIII)	This work
pAB9	**pH118::TTP128**. Hyg^R. Insertion of the *hph5* cassette interrupting *TTP128* gene (EcoRI/HindIII)	This work
pAB10	**pH118::TTP208**. Hyg^R. Insertion of the *hph5* cassette interrupting *TTP208* gene (EcoRI/HindIII)	This work
pAB12	**pH118::TTP081**. Hyg^R. Insertion of the *hph5* cassette interrupting *TTP081* gene (EcoRI/HindIII)	This work

pAB13	**pH118::TTP085**. Hyg^{R}. Insertion of the *hph5* cassette interrupting *TTP085* gene (EcoRI/HindIII)	This work
pAB14	**pH118::TTP140.** Hyg^{R}. Insertion of the *hph5* cassette interrupting *TTP140* gene (EcoRI/HindIII)	This work
pAB15	**pk18::TTP140**. Kan^{R}. Insertion of the *kat* cassette interrupting *TTP085* gene (EcoRI/HindIII)	This work
pAB16	**pK18::TTP211**. Kan^{R}. Insertion of the *kat* cassette interrupting *TTP211* gene (EcoRI/HindIII)	This work
pAB17	**pH118::TTP211.** Hyg^{R}. Insertion of the *hph5* cassette interrupting *TTP085* gene (EcoRI/HindIII)	This work
pAB19	**pH118::TTP191**. Hyg^{R}. Insertion of the *hph5* cassette interrupting *TTP191* gene (BglII)	This work
pAB22	**pUC19::ΔpilA4:kat.** Kan^{R}. *pilA4* deletion, replaced by the *kat* cassette, flanked by upstream and downstream sequences of *pilA4* gene (EcoRI/HindIII)	This work
pAB25	**pUC119::ΔrepA2::kat**. Amp^{R},Kan^{R}. Deletion of *TTP145* gene, replaced by the *kat* cassette, flanked by upstream and downstream sequences of the deleted gene (EcoRI/HindIII)	This work
pAB26	**pH118::TTP145**. Hyg^{R}. Insertion of the *hph5* cassette interrupting *TTP145* gene (EcoRI/HindIII)	This work
pAB32	**pUC19:TTP211::kat**. Amp^{R}, Kan^{R}. Deletion of *TTP211* gene, replaced by the *kat* cassette, flanked by upstream and downstream sequences (EcoRI/HindIII)	This work
pAB36	**pUC19::ΔpilQ::kat.** Kan^{R}. *pilQ* deletion, replaced by the *kat* cassette, flanked by upstream and downstream sequences of *pilQ* gene (EcoRI/HindIII)	This work
pAB38	**pH118::ΔpilQ::kat**. Hyg^{R}, Kan^{R}. pH118 derivative harbouring flanking regions of *pilQ* gene, interpersed by *kat* cassette (XbaI)	This work
pAB43	**pUC19::ΔpilQ::hyg**. Amp^{R}, Hyg^{R}. *pilQ* deletion, replaced by the *hph5* cassette, flanked by upstream and downstream sequences of *pilQ* gene (BglII)	This work
pAB48	**pUC19:ΔrecA:kat.** Amp^{R}, Kan^{R}. pUC19 derivative harbouring up and downstream sequences of *recA* gene, replaced by the *kat* cassette (EcoRI/HindIII)	This work
pAB51	**pUC19:: ΔpilQ::hyg**. Amp^{R}, Hyg^{R}. Substitution plasmid. pUC19 derivative, enclosing up and dowsntream flanking regions of *pilQ* gene, which is replaced by the *hph5* cassette (XbaI)	This work
pAB52	**pUC19:: ΔpilA4::hyg**. Amp^{R}, Hyg^{R}. pUC19 derivative, enclosing up and dowsntream flanking regions of *pilA4* gene, which is replaced by the *hph5* cassette (XbaI)	This work
pAB54	**pUC19::ΔpulE::kat.** Amp^{R}, Kan^{R}. pUC19 derivative, enclosing up and dowsntream flanking regions of *pulE* gene (*TTC1844*), which is replaced by the *kat* cassette (XbaI)	This work
pAB55	**pUC19::ΔpulE::hyg**. Amp^{R}, Hyg^{R}. pUC19 derivative, enclosing up and dowsntream flanking regions of *pulE* gene (*TTC1844*), which is replaced by the *hph5* cassette (XbaI)	This work
pAB59	**pUC19::ΔTTC1415::H.** Amp^{R}, Hyg^{R}. pUC19 derivative, enclosing up and dowsntream flanking regions of *pilT2* gene, which is replaced by the *hph5* cassette (EcoRI/PstI)	This work
pAB66	**pUC19::Δ1429-30::kat.** Amp^{R}, Kan^{R}. pUC19 derivative, enclosing up and dowsntream flanking regions of *TTC1429* (*nurA1*) and *TTC1430* (*cptB*) genes. The deleted sequence is replaced by the *kat* cassette (XbaI)	This work
pAB67	**puc19::Δ1429-30::hyg.** Amp^{R}, Hyg^{R}. pUC19 derivative, enclosing up and dowsntream flanking regions of *TTC1429* (*nurA1*) and *TTC1430* (*cptB*) genes. The deleted sequence is replaced by the *hph5* cassette (XbaI)	This work
pAB69	**pH118:Δ18xx**. Hyg^{R}. pH118 derivative, enclosing upstream sequence of *TTC1836* and dowsntream sequence of *TTC1847* genes. Substitution of *TTC18XX* cluster (EcoRI/HindIII)	This work
pAB70	**pk118::cptA**. Kan^{R}. Insertion of the *kat* cassette interrupting *TTC1879* gene (*cptA*) (HindIII/BamHI)	This work
pAB71	**pH118::cptA.** Hyg^{R}. Insertion of the *hph5* cassette interrupting *TTC1879* gene (*cptA)* (HindIII/BamHI)	This work
pAB83	**pUC18::ΔnurA1::kat.** Amp^{R}, Kan^{R}. pUC18 derivative vector replacing *nurA (TTC1429)* by *kat* cassette, flanked by up and dowsntream regions. (EcoRI/HindIII)	This work
pAB84	**pUC18::ΔnurA1::hyg.** Amp^{R}, Hyg^{R}. pUC18 derivative vector replacing *nurA*	This work

	(TTC1429) by *hph5* cassette, flanked by up and dowsntream sequences (EcoRI/HindIII)	
pAB101	**pK18::TTC0656.** Kan^{R}. Insertion of the *kat* cassette interrupting *TTC0656* gene (*primPol*) (EcoRI/HindIII)	This work
pAB102	**pH118::TTC0656**. Hyg^{R}. Insertion of the *hph5* cassette interrupting *TTC0656* gene (*primPol*) (EcoRI/HindIII)	This work
pAB106	**pUC19::ΔcptA::kat**. Amp^{R}, Kan^{R}. pUC19 derivative where *cptA* gene is replaced by the *kat* cassette, flanked by upstream and dowsnream regions of *TTC1879* (HindIII/BamHI)	This work
pAB107	**pUC19::ΔcptA::hyg**. Amp^{R}, Hyg^{R}. pUC19 derivative where *cptA* gene is replaced by the *hph5* cassette, flanked by upstream and dowsnream regions of *TTC1879* (HindIII/BamHI)	This work
pAB109	**pmH::c1879**. Hyg^{R}. Complementation. pMH185 derivative expressing *cptA* gene (BamHI/HindIII)	This work
pAB110	**pMK::c1879**. Kan^{R}. Complementation. pMK184 derivative expressing *cptA* gene (BamHI/HindIII)	This work
pAB112	**pmH::c1430.** Hyg^{R}. Complementation. pMH185 derivative expressing *cptB* gene (*TTC1430*) (BamHI/HindIII)	This work
pAB113	**pMK::c1430**. Kan^{R}. Complementation. pMK184 derivative expressing *cptB* gene (*TTC1430*) (BamHI/HindIII)	This work
pAB114	**pmH::c1844**. Hyg^{R}. Complementation. pMH185 derivative expressing *pulE* gene (*TTC1844*) (EcoRI/HindIII).	This work
pAB115	**pMK::c1844**. Kan^{R}. Complementation. pMK184 derivative expressing *pulE* gene (*TTC1844*) (EcoRI/HindIII)	This work
pAB122	**pUC18::cptB::kat**. Amp^{R}, Kan^{R}. pUC18 derivative where *cptB* gene (*TTC*1430) is replaced by the *kat* cassette, flanked by upstream and dowsnream regions of *TTC1430* (HindIII/EcoRI)	This work
pAB131	**pSW23::ISBstI::KATrpmB::Pnqo::Tnp.** Cm^{R}, Kan^{R}. Background plasmid for transposon mutagenesis. Transposase from *Geobacilllus stearothermophilus T-6* is under the control of *T. thermophilus* Pnqo promoter expressing Kan resistance which is controllled by *T. thermophilus rpmB* promoter and terminator, flanked by the IS left and right overhangs of such transposase (PstI/NotI)	This work
pAB132	**pSW23::ISNth2::KATrpmB::Pnqo::Tnp**. Cm^{R}, Kan^{R}. Background plasmid for transposon mutagenesis. Transposase from *Natraerobius thermophilus* JW/NM-WLN-LF is under the control of *T. thermophilus* Pnqo promoter expressing Kan resistance, which is controlled by *T. thermophilus rpmB* promoter and terminator, flanked by the IS left and right overhangs of such transposase (PstI/NotI)	This work
pAB134	**pET11a::TTHA0522**. Amp^{R}. Expression of *T. thermophilus* HB8 *TTHA0522* gene (ttHerA homolog) (EcoRI/NdeI)	This work
pAB135	**pET28b::TTHA0522** .Kan^{R}. Expression of *T. thermophilus* HB8 *TTHA0522* gene (ttHerA homolog) (EcoRI/NdeI)	This work
pAB140	**pMK184::TTHA0522.** Expression of *T. thermophilus* HB8 *TTHA0522* gene. (EcoRI/NdeI)	This work
pAB145	**pUC18::Δ*primpol::kat.*** Amp^{R}, Kan^{R}. pUC18 derivative where *primPol* gene (*TTC0656*) is replaced by the *kat* cassette, flanked by upstream and downstream regions of *TTC0656* (PstI/EcoRI)	This work
pAB146	**pUC18:: Δ*primpol::hyg*.** Amp^{R}, Hyg^{R}. pUC18 derivative where *primPol* gene (*TTC0656*) is replaced by the *hph5* cassette, flanked by upstream and downstream regions of *TTC0656* (PstI/EcoRI)	This work
pAB148	**pH118::TTC1450**. Hyg^{R}. Insertion of the *hph5* cassette interrupting *TTC1450* gene (EcoRI/HindIII)	This work
pAB149	**pK18::TTC1450**. Kan^{R}. Insertion of the *kat* cassette interrupting *TTC1450* gene (EcoRI/HindIII)	This work
pAB150	**pH118::herA**. Hyg^{R}. Insertion of the *hph5* cassette interrupting *TTC0147* gene (*ttherA*) (EcoRI/PstI)	This work
pAB151	**pK18::herA**. Kan^{R}. Insertion of the *kat* cassette interrupting *TTC0147* gene (*ttherA*) (EcoRI/PstI)	This work
pAB152	**pH118::*pyrE*.** Hyg^{R}. Insertion of the *hph5* cassette interrupting *TTC1380* gene (*pyrE*) (EcoRI/HindIII)	This work
pAB153	**pK18::*pyrE***. Kan^{R}. Insertion of the *kat* cassette interrupting *ttc TTC1380* gene (*pyrE*) (EcoRI/HindIII)	This work
pAB160	**pUC18::Δ*cptB::hyg***. Amp^{R}, Kan^{R}. pUC18 derivative where *cptB* gene (*TTC1430*) is replaced by the *hph5* cassette, flanked by upstream and dowsnream regions of *cptB* (HindIII/EcoRI)	This work

pAB175	**pMH::ttherA**::**sYFP.** Hyg^R. Expression plasmid, pMHPnqosYFP derivative where the gene *herA* (*TTC0147*) is fused to sYFP (BcuI)	This work
pAB176	**pMH:: ttherA::sGFP.** Hyg^R. Expression plasmid, pMHnqosGFP derivative where the gene *herA (TTC0147)* is fused to sGFP (BcuI)	This work
pAB181	**pK18::TTC0474**. Kan^R. Insertion of the *kat* cassette interrupting *TTC0474* gene (*ftsk*) (EcoRI/HindIII)	This work
pAB182	**pH118::TTC0474**. Hyg^R. Insertion of the *hph5* cassette interrupting *TTC0474* gene (*ftsk*) (EcoRI/HindIII)	This work
pAB183	**pK18::TTC1026**. Kan^R. Insertion of the *kat* cassette interrupting *TTC1026* gene (EcoRI/PstI)	This work
pAB184	**pH118::TTC1026.** HygR. Insertion of the *hph5* cassette interrupting *TTC1026* gene (EcoRI/PstI)	This work
pAB185	**pK18::TTC1836.** Kan^R. Insertion of the *kat* cassette interrupting *TTC1836* gene (*comGC*-like) (EcoRI/HindIII)	This work
pAB186	**pH118::TTC1836.** Hyg^R. Insertion of the *hph5* cassette interrupting *TTC1836* gene (*comGC*-like) (EcoRI/HindIII)	This work
pAB187	**pK18::TTC1839.** Kan^R. Insertion of the *kat* cassette interrupting *TTC1839* gene (*comGA*-like) (EcoRI/HindIII)	This work
pAB188	**pH118::TTC1839**. Hyg^R. Insertion of the *hph5* cassette interrupting *TTC1839* gene (*comGA*-like) (EcoRI/HindIII)	This work
pAB199	**pZERO::Nth2 tnpase.** Amp^R**.** pZEROBluntTOPO commercial vectors (Invitrogen) harbouring the synonimous synthesized transposase *ISNth2* from *Natraerobius thermophilus*	This work
pAB200	**pZERO::BstI tnpase.** Amp^R**.** pZEROBluntTOPO commercial vectors (Invitrogen) harbouring the synthesized transposase *ISBstI* from *Geobacilllus stearothermophilus* T-6	This work
pAB201	**pET28b::cptA.** Kan^R. Expression of *T. thermophilus* HB27 *TTC1879* gene (*cptA*) (NdeI//HindIII)	This work
pAB207	**pUC18::ΔherA::kat.** Amp^R, Kan^R. pUC18 derivative, enclosing up and dowsntream flanking regions of *ttherA* gene (*TTC0147*), which is replaced by the *kat* cassette (XbaI)	This work
pAB208	**pUC18::ΔherA::hyg.** Amp^R, Hyg^R. pUC18 derivative, enclosing up and dowsntream flanking regions of *ttherA* gene (*TTC0147*), which is replaced by the *hyg* cassette (XbaI)	This work
pAB209	**pK::herA::sYFP.** Kan^R. Recombinant plasmid, pK118 derivative where TtHerA (*TTC0147*) is fused to sYFP (BcuI)	This work
pAB210	**pK:cptA::sYFP.** Kan^R. Recombinant plasmid, pK118 derivative where the C-terminal of ttCptA (*TTC1879*) is fused to sYFP (BcuI)	This work
pAB213	**pH:cptA::sYFP**. Hyg^R. Recombinant plasmid, pH118 derivative where the C-terminal of ttCptA (*TTC1879*) is fused to sYFP (BcuI)	This work
pAB214	**pH::Argo::sGFP**. Hyg^R. Recombinant plasmid, pK118 derivative where the C-terminal of TtAgo (*TTP026*) is fused to sGFP (BcuI)	This work
pAB219	**pH::herA::sYFP.** Hyg^R. Recombinant plasmid, pK118 derivative where the C-terminal of TtHerA (*TTC0147*) is fused to sYFP (BcuI)	This work
pAB221	**pH::Argo::sYFP.** Hyg^R. Recombinant plasmid, pK118 derivative where the C-terminal of ttAgo (*TTP026*) is fused to sYFP (BcuI)	This work
pAB224	**pMH::cptA::sYFP**. Hyg^R. Expression plasmid, pMHPnqosYFP derivative where TtCptA is fused to sYFP (BcuI)	This work
pAB225	**pMH::cptA::sGFP**. Hyg^R. Expression plasmid, pMHPnqosGFP derivative where TtCptA is fused to sGFP (BcuI)	This work

ANNEX II

TABLE II. Oligonucleotides employed in this dissertation. The sequence that hybridizes is indicated In capital letters, whereas the restriction site is underlined.

Name	Use	Sequence 5'-->3'
kat1	Reverse. Upstream amplification of *kat* gene. Sequencing	CCTTTTTCCCCGCATCC
kat2	Forward. Downstream amplification of *kat* gene. Sequencing	GAAACTTCTGGAATCGC
kat3	Forward. *kat* cassette amplification and sequencing	GGAACGAATATTGGATA
kat4	Reverse. *kat* cassette amplification and sequencing	AGAAATTCTCTAGCGAT
M13F	Forward. Sequencing of pUC, pK, pH, pMK, pMH, pWUR derivatives	GTAAAACGACGGCCAGT
M13R*	Reverse. Sequencing of pUC, pK, pH, pMK, pMH derivatives	CAGGAAACAGCTATGAC
F24	Forward. Sequencing of pUC, pK, pH, pMK, pMH, pWUR derivatives	CGCCAGGGTTTTCCCAGTCACGAC
R24	Reverse. Sequencing of pUC, pK, pH, pMK, pMH derivatives	AGCGGATAACAATTTCACACAGGA
T7prom	Forward. Sequencing pET derivatives	CGACTCACTATAGGGGAATTG
T7term	Reverse. Sequencing pET derivatives	GCTAGTTATTGCTCAGCGG
TTC0313 dir	Forward. Check *TTC0313* mutation	CTTTACGAGGCCCTCTTGGAG
TTC0313 rev	Reverse. Check *TTC0313* mutation	CCACCGCTCGGGGAC
CEC10	Reverse. Construction of pCBG1 and pCBG2	aaaagaacgttttcCTCGGGTCCGTGGCCTC
CEC11	Forward. Construction of pCBG1 and pCBG2	aaaaaagcttGCCTGAGGGCCAT
CEC61	Forward. Construction of pCBG3	atcaagcttGAGGCAAGGCGGTCTAAGCTTAG
CEC62	Reverse. Construction of pCBG3	tcagaattcTCCGCTCACGCTCCGTG
CEC78	Forward. Construction of pCEC30 and pCEC32	ttcgaattcGTCGCTGGTCATGGCGTC
CEC79	Reverse. Construction of pCEC30 and pCEC32	atcaagcttCACCGCTACCTGGTGGACTC
CEC80	Forward. Construction of pCEC31	ttcgaattcCTACTACTTCGCCCTGGCCATC
CEC81	Reverse. Construction of pCEC31	atcaagcttGGTTCTCCCCGCTCACGTAG
CEC84	Forward. Construction of pCEC33	ttcgaattcGTGGCGATGAGGATCTCCAG
CEC85	Reverse. Construction of pCEC33	atcaagcttGTCCGGATAGACGGCAAGCTC
CEC86	Forward. Construction of pCEC37	ttcgaattcGTGAGCGAGCGACAGGACAC
CEC87	Reverse. Construction of pCEC37	atcaagcttGTCTCGTTCCAGAGCATCGTG
CEC88	Forward. Construction of pCEC34	ttcgaattcGAGATCCTGGAGGTCCTGGA

		C
CEC89	Reverse. Construction of pCEC34	atcaagcttCACGGTGCCGATGAGGGTG
CEC92	Forward. Construction of pCEC35	atcaagcttCCTCTCCAACCCCGACATG
CEC93	Reverse. Construction of pCEC35	ttcgaattcACTCGTCCATGGGGTCCTCTC
CEC94	Forward. Construction of pCEC36	ttcgaattcGCCAAAAGCCGCTCCTTCTC
CEC95	Reverse. Construction of pCEC36	atcaagcttCGGGACGAGGTCTTTCTTTC
CEC99	Forward. Detect NCE	ATCGCCTTTTCCCTGGAGC
CEC100	Reverse. Detect NCE (NAR I specific)	TGGATCAGGTGGGAGAGTCG
CEC101	Reverse. Detect no NCE (HB27 specific)	CACCTGCGATGTAGAGGGTTC
CEC102	Reverse. Construction of pCEC40	atcaagcttACGGAAGCGCAGCTTAGGGA G
CEC103	Forward. Construction of pCEC40	ttcgaattcAAGAAGGTCTCGCCGGACTG
AB51	Forward. Construction of pAB22 and pAB52. Upstream amplification of pilA4	atcaagcttTGCTGAAGCTTGGCGGCAAC
AB52	Reverse. Construction of pAB22 and pAB52. Upstream amplification of *pilA4*	aaaatctagaCAGCCTGGTTGCTGGTGTA GGAG
AB53	Forward. Construction of pAB22 and pAB52. Downstream amplification of *pilA4* gene	aaaaatctagagTTAAAGTATGGCCTCGG CGCTAG
AB54	Reverse. Construction of pAB22 and pAB52. Downstream amplification of *pilA4* gene	aaaagaattcGGGAGTTAGGCTTGGGATT GTG
AB63	Forward. Construction of pAB3	ttcgaattcCTTCATCCCCACCAAGTTTGA C
AB64	Reverse. Construction of pAB3	aacaagcttGTCCTTCACCTTCTTGAGCT CCAG
AB65	Forward. Construction of pAB4 and pAB9	tcagaattcGACCACGAGGTCTTCCAGCT C
AB66	Reverse. Construction of pAB4 and pAB9	actaagcttAGTGGAGGAGGACGGCGAT C
AB67	Forward. Construction of pAB5and pAB10	ttcgaattcTCCTCAAGGAGGCCCTCTG
AB68	Reverse. Construction of pAB5 and pAB10	atcaagcttGAAGTCCGCGAACTCGGTAA G
AB69	Forward. Construction of pAB6 and pAB19	ttcgaattcGAGCGGCAAGACCACCACCC TG
AB70	Reverse. Construction of pAB6 and pAB19	atcaagcttGTGAGGTTGCGCACGAAG
AB71	Forward. Construction of pAB7 and pAB12	ttcgaattcTTGCGAAGCGGCCTTTAC
AB72	Reverse. Construction of pAB7 and pAB12	atcaagcttCTTGAGGTCCTTCAGCGCCT C
AB73	Forward. Construction of pAB8 and pAB13	aaaagaattcGCGCATCTTGCGGAAGCGT TC
AB75	Reverse. Construction of pAB8 and pAB13	aaataagcttTCCCTGCAGAACCGCTACA C
AB76	Forward. Construction of pAB14 and pAB15	aaaagaattcCTCAATGCGAGCGGGTTCT C
AB77	Reverse. Construction of pAB14 and pAB15	aaaaaagcttGTCCGTGGGCAGAACCTTG
AB80	Forward. Construction of pAB16 and pAB17	atcaagcttCTCCCCCACTTCCAG
AB81	Reverse. Construction of pAB16 and pAB17	tcagaattcTCATCCTGGACGAGGTGCAG

AB82	Forward. Internal primer; check deletion of *pilA4*	aaaCTGCATGATCGCAGGCCAC
AB83	Forward. Amplification and check of *pilA4*	tcaCTGAACGACGCGGTGAAGTTC
AB84	Reverse. Amplification and check of *pilA4*	ttcGAAGACGGCCAAAAGGTGTTG
AB86	Forward. Construction of pAB25. Upstream amplification of *TTP145*	aaagaattcATGACCTCCGCGTCCAGAG
AB87	Reverse. Construction of pAB25. Upstream amplification of *TTP145*	aaatctagaGGAGGGCCTTACCGGATTACTG
AB88	Reverse. Amplification of *TTP145-TTP146* region	AAGCTTAGACCGCCTTGCCTC
AB89	Forward. Amplification of *TTP145-TTP146* region	gaattcAGTCCACCACCACCACCAC
AB90	Forward. Construction of pAB25. Downstream amplification of *TTP145*	aaatctagaTGCTGCTCGGATGGTTTG
AB91	Reverse. Construction of pAB25. Downstream amplification of *TTP145*	atcaagcttCCTCCAGGGAACATCCAGTAGAG
AB92	Forward. Check *pilA4* mutation	aaaTGCTGAAGCTTGGCGGCAAC
AB93	Reverse. Check *pilA4* mutation	aaaAGAATTCGGGAGTTAGGCTTGGGATTGTG
AB94	Forward. Deletion of *TTP211*	tcagaattcCAGGGCCTCCATGAGCTTG
AB95	Reverse. Deletion of *TTP211*	atcaagcttGTGGATGCGGTCTTTGACGAC
AB100	Reverse. Construction of pAB36. Upstream amplification of *pilQ*	aaatctagaCGAGGTAGCCCTCCTTGCTC
AB101	Forward. Construction of pAB36. Upstream amplification of *pilQ*	aaagaattcCTGGAGGACCTCTCCCGCTTCAG
AB102	Reverse. Construction of pAB36. Downstream amplification of *pilQ*	atcaagcttGCACCGCTTCCTTCTCCATC
AB103	Forward. Construction of pAB36. Downstream amplification of *pilQ*	tcatctagaGACTTCAAGAACTGGGCCGAG
AB107	Forward. Construction of pAB51. Upstream amplification of *pilQ*	aaagaattcCACCTTGACCGTCTACGTGCTG
AB108	Reverse. Construction of pAB51. Upstream amplification of *pilQ*	atcagatctGAGGCCCTTTTCCGCCAC
AB109	Forward. Construction of pAB51. Downstream amplification of *pilQ*	atcagatctCAATGTCCCCATGCGGTTTC
AB110	Reverse. Construction of pAB51. Downstream amplification of *pilQ*	ttcaagcttGTAGATGGCGTCGTGGACCTC
AB111	Forward. Internal primer; check deletion of *recA*	CACGCTGGCGTAGAACTTCAG
AB112	Forward. Check *recA* mutation	GACCAGGAGCAGAAGAACCAGG
AB113	Reverse. Check *recA* mutation	GTCCTCTTCGGCCTGGAGTG
AB114	Forward. Internal primer; check *pilQ* deletion	GAGGAGATCGTTCTGGACCCAG
AB115	Forward. Check *pilQ* mutation	CACGTCCCTCAGGTCCTTGTG
AB116	Reverse. Check *pilQ* mutation	CTTCCCAAGAGGAGGCCCAG
AB117	Reverse. Insertion of *BglII* site upstream *pilA4*	atcagatctCTCCCCGTGCCTTTCTTAC
AB118	Forward. Insertion of *BglII* site downstream *pilA4*	atcagatctCTCCTCCTCTGAGCCCACTCAC
AB119	Reverse. Insertion of *BglII* site upstream *pulE*	atcagatctGATGCTTCCCCGCTCCAAG
AB120	Forward. Check *pulE* mutation	atcagatctCTACGAGGCCACGGACGAG

AB122	Forward. Internal primer; check *pulE* deletion	CAGCGACCTCCACCTGGAC
AB123	Forward. Amplification *pulE*; check *pulE* mutation	atcaagcttCAGGTCCCCTCTTGGACCTCAG
AB124	Reverse. Amplification *pulE*; check *pulE* mutation	ttcgaattcAGGAGGCGCCGATCAGGAAG
AB126	Forward. Construction of pAB48. Upstream amplification of *recA*	aaaaaatctagaGAGAAGGCGGCCGAGG
AB127	Reverse. Construction of pAB48. Upstream amplification of recA	aaaaaaaagcttGTCCGGGGTGTCGTCAATG
AB128	Forward. Construction of pAB48. Downstream amplification of *recA*	aaaaaatctagaCTCAATCGCCTTCAGGGCG
AB129	Reverse. Construction of pAB48. Downstream amplification of *recA*	aaaaaagaattcGGAGATGACCCTCGCCG
AB130	Forward. Check *recA* mutation	ACAAGCCCTTCAAGCCCCA
AB131	Reverse. Check *recA* mutation	TTGAGCTCCTCCTGGAGGGA
AB133	Reverse. Construction of pAB69. Upstream amplification *TTC1836*	atccctgcagTAGACCTCGTCCTCAGGGTCCTC
AB134	Forward. Construction of pAB69. Downstream amplification *TTC1847*	atacaagcttGCTCCAGACCCGCTCCACCCGGGCCTGG
AB135	Reverse. Construction of pAB69. Downstream amplification *TTC1847*	attagaattcGCCACCTTCCGGGACCTCACGGGGAAGAG
AB136	Forward .Construction of pAB69. Upstream amplification *TTC1836*	agtctctagaCTCTGCACCTCCCTAACCGGGGCTTCCCTTC
AB137	Forward. Check *pilT* mutation	GGGACTCCGAGACGGCCAAGATCGCCACC
AB138	Reverse. Check mutation *pilF* and *pilT*	GGTAGTCGGCCTCTTCCTCCTTGGTGCGCATGGG
AB139	Forward. Check mutation *pilF* and *pilT*	ATTAAAGCTTCCCGCCTTAGGATCTCCGTCTGGCGG
AB140	Forward. Check *pilT2* mutation	TAGCTTCTCCTGCCGCTTCTTCTTGCCCAC
AB141	Reverse. Check *pilT2* mutation	ATGGCCTGGGTGAACTCCATGTTGTTGGTG
AB142	Forward. Check mutant *pilF::kat*	CTCTTCACCTTCCGCGAGGAC
AB143	Reverse. Check mutant *pilF::kat*	GAGGACGCTGCGGAGCGCCTTG
AB145	Forward. Check *pilQ* deletion	GGGCTCCTCTGGTACTACCTGCTCATCGTG
AB146	Reverse. Check *pilQ* deletion	GTAGCGCTCCACCCGTTCCTGGATCTCCTC
AB147	Forward. Check *pulE* deletion	GTAGAAGGCGCCCAGGAAGAAGAGGAGGATCC
AB148	Reverse. Check *pulE* deletion	CTCTTCGGCCACCCGGACCCCGTGGAGGAG
AB149	Forward. Check *cptB* deletion	GAGTACTCCCGCGAAGCCACATCGAAGGCCGAGTG
AB150	Reverse. Check *cptB* deletion	ATGTCGGAGCAAGTGGTGGACATCGATCGG
AB151	Forward. Construction of pAB66 and pAB67. Upstream amplification of *nurA1*	tcagaattcCAGCCTACGCCAACTGCAC
AB152	Reverse. Construction of pAB66 and pAB67. Upstream amplification of *nurA1*	aaatctagaGTGCTACTCGGGGAAGTAACCATC
AB153	Forward. Construction of pAB66 and pAB67.	aaatctagaGTACGACCTGGGCCACCTG

	Downstream amplification of *cptB*	
AB154	Reverse. Construction of pAB66 and pAB67. Downstream amplification of *cptB*	atcaagcttGAGTCGAGACCGCGATCAAG
AB155	Forward. Check double mutants *nurA1-cptB*	GACGGATGCAGTGCGGAAG
AB156	Reverse Check double mutants *nurA1-cptB*	CTACGCCTTCGTCTACCTGGAC
AB159	Reverse. Amplification of Bleomycin cassette	aaactgcagCTTCCGGCTCGTATGTTGTG TGG
AB160	Forward. Amplification of Bleomycin cassette	agatatcGGCGGCGCAGGCCTTCCTGG
AB161	Forward. Amplification of Hygromicin cassette	agccggagtctagaCCCGGGAGTATAAC
AB162	Reverse. Amplification of Hygromycin cassette	aggcggctgcagACGTTCAAAATGGTATG CGTTTTGACAC
AB166	Forward. Amplification of Kanamycin cassette	CAGGTGGATACTTAGAGAAAGTGTAC AAACTG
AB168	Reverse. Amplification of Kanamycin cassette	CCGTGTCGTTCTGTTCCACTCCTGAA TCCC
AB169	Forward. Construction of pAB70 and pAB71	atcaagcttGAGTTATTGGCCGCGCTTC
AB170	Reverse. Construction of pAB70 and pAB71	aaaccatggCATGCGGGTGCTCAGGTG
AB171	Forward. Check *TtAgo* deletion	GTCTTCCTTCTCCCTCCGGAC
AB172	Reverse. Check *TtAgo* deletion	CTTCGGGCTTTCCCTGGAG
AB173	Forward. Construction of pAB69. Downstream amplification *TTC1847*	aaatctagaGTAGAGGGCCTTTTCCTCCG CCTCCGTG
AB174	Reverse. Construction of pAB70 and pAB71. *BamHI* site	aaaggatccCATGCGGGTGCTCAGGTG
AB175	Reverse. Construction of pAB106 and pAB107. Upstream amplification of *cptA*	aaatctagaGACGGCCTCCAGATAGCCA TC
AB176	Forward. Construction of pAB106 and pAB107. Upstream amplification of *cptA*	atcaagcttCATAGACGAGCAGGTGCTCC AG
AB177	Forward. Construction of pAB106 and pAB107. Downstream amplification of *cptA*	aaatctagaCAGGGACCGCTGTTCTACGT G
AB178	Reverse. Construction of pAB106 and pAB107. Downstream amplification of *cptA*	aaagaattcCTTGCAGGCCGCTATCACAC
AB179	Forward. Construction of pAB122 and pAB160. Upstream amplification of *cptB*	aaagaattcGAGCAGTTGCTGGAGAGCC TG
AB180	Reverse. Construction of pAB122 and pAB160. Upstream amplification of *cptB*	aaatctagaCTGGCGCAGGGCAGGTAC
AB181	Forward. Construction of pAB83 and pAB84. Upstream amplification of *nurA1*	aaatctagaGTACCTGCCCTGCGCCAG
AB182	Forward. Construction of pAB106 and pAB107. Upstream amplification of *cptA*	AAACTGCAGcatagacgagcaggtgctccag
AB183	Reverse. Construction of pAB101 and pAB102	aataagcttGACCTCCTCCGCTTCCCTG
AB184	Forward. Construction of pAB101 and pAB102	aaagaattcCACCCTCGAGGCCTGGTG
AB185	Reverse. Amplification of *cptB* for complementation	atcaagcttCAGGTGGCCCAGGTCGTAC
AB186	Forward. Amplification of *cptB* for complementation	aaagaatccGAGATGAGCCGCACCGTG
AB187	Reverse. Amplification of *pulE* for complementation	aaagaattcGAAGGCCCTGGACCCCTAC
AB188	Reverse. Amplification of *pulE* for complementation	atcaagcttCTCCAGGGAGAGGGTGTGG AG

AB189	Reverse. Amplification of *cptA* for complementation	atcaagcttGATCCTTTGCGGGCCGATAC
AB190	Forward. Amplification of *cptA* for complementation	aaaggatccGACGACGGACGGATCCACATC
AB191	Reverse. Check *cptA* deletion	CTGCTGGAGTGAAGGGGGTG
AB192	Forward. Check *cptA* deletion	GATGTGCGCCGCTGGTATG
AB193	Forward. Check *cptA* mutation	CAGGCCATTCCGGACCTG
AB194	Reverse. Check *cptA* mutation	CAACGTGCTGCTTGGCCTG
AB195	Forward. Construction of pAB145 and pAB146. Upstream amplification of *primPol*	aaagaattcCTCTGGAAAGCTGGGCTCG
AB196	Reverse. Construction of pAB145 and pAB146. Upstream amplification of *primPol*	aaatctagaGACCACAACGGCACGCAC
AB197	Forward. Construction of pAB145 and pAB146. Downstream amplification of *primPol*	aaatctagaCTGGCCCTCAAGCGGAAG
AB198	Reverse. Construction of pAB145 and pAB146. Downstream amplification of *primPol*	aaactgcagGAAGCCGGGTCAGGTCCTC
AB199	Forward. Amplification of Pnqo promoter	aaaagaattcCCCCTCTAGACCTCCCTCCAG
AB200	Reverse. Amplification of Pnqo promoter	aaaactgcagCAACGGTCGACATATGCCCCTC
AB201	Forward. Construction of pAB131 and pAB132. Amplification of 3' of Pnqo for Gibson assembly with IS	aaaaCACGAAAGGAGGGGCATATGATG
AB202	Reverse. Construction of pAB131 and pAB132. Amplification of 3' of Pnqo for Gibson assembly with IS	aaaactagtgGATCCCCCGGGCTGCAGGAATTC
AB205	Forward. Construction of pAB131. Amplification of *kat* cassette flanked by left IR from *ISBstI* (Tnpase from *Bacillus staerothermophilus*)	aaatctagaTGTTTCTCCAACGTCCGGATCCTGATTTTACATACATATCCTAAATTGGAAAGGCGAGGAGGAG
AB206	Reverse. Construction of pAB131. Amplification of *kat* cassette flanked by right IR from *ISBstI* (Tnpase from *Bacillus staerothermophilus*)	aaagcggccgcTGTTTCTTTCGCAAACGGCTCCCGCGAAAGGGAGTAAAACGATTTATACAAAGCTTAGGCAGCGCAAC
AB207	Forward. Construction of pAB132. Amplification of *kat* cassette flanked by left IR from *ISNth2* (Tnpase from *Natraerobius thermophilus*)	aaatctagaGTTATTGTAAAATTACTGTAGGATTTTTTGGAGAAAGCAGGAAGGAAAGGAAGGCGAGGAGGAG
AB208	Reverse. Construction of pAB132. Amplification of kat cassette flanked by right IR from *ISNth2* (Tnpase from *Natraerobius thermophilus*)	aaagcggccgcGATCGTGTCAAGCCACTGTAGGTTAAAAAGAAACCTTAATTCTCCTAGATAAGCTTAGGCAGCGCAACCTG
AB209	Reverse. Check plasmid pNCY::*rpmB*::*kat*	GTTTCCTCTCCGCCCAGGTC
AB210	Forward. Check plasmid pNCY::*rpmB*::*kat*	GACGGTGAGAGGTGGGAGAAG
AB211	Forward. Sequencing pSW23, pAB131, pAB132 and derivatives	CCGTCACAGGTATTTATTCGGCG
AB212	Reverse. Sequencing pSW23, pAB131, pAB132 and derivatives	CCTCACTAAAGGGAACAAAAGCTG
AB213	Forward. Sequencing pSW23, pAB131, pAB132 and derivatives	GATATCGCGCGCGTAATACGAC
AB214	Forward. Sequencing pSW23, pAB131, pAB132 and derivatives	CTGCGAAGTGATCTTCCGTCAC
AB215	Reverse. Sequencing pSW23, pAB131, pAB132 and derivatives	CACAGGAACACTTAACGGCTGAC
AB216	Reverse. Sequencing pSW23, pAB131, pAB132 and derivatives	CAAGCCAGGGATGTAACGCAC

AB217	Reverse. Check and sequencing of pAB131, pAB132 and derivatives from inside *kat* cassette	CACCTTCCACTCACCGGTTG
AB218	Forward. Check and sequencing of pAB131, pAB132 and derivatives from inside *kat* cassette	CAACCGGTGAGTGGAAGGTG
AB219	Reverse. Amplification and check of *TTHA0522*	CTACCTGAAGAACTCCCGGCGCAG
AB220	Forward. Amplification and check of *TTHA0522*	GTGAAGCGTATCGGCGTGGTCTTG
AB221	Forward. Construction of pAB150 and pAB151	aaagaattcGAGGTGGCCTACCTCAACCTG
AB222	Reverse. Construction of pAB150 and pAB151	aaactgcagGAGCTCGTCCAGGACGATGAAG
AB223	Forward. Construction of pAB148 and pAB149	aaagaattcGATGCGGTGGGATAGGGAGAG
AB224	Reverse. Construction of pAB148 and pAB149	atcaagcttGTAAATTGCCTTGTACCCCGGAG
AB225	Forward. Construction of pAB152 and pAB153	aaagaattcCTAGACCTCCTCCAAGGGCAC
AB226	Reverse. Construction of pAB152 and pAB153	atcaagcttATGGACGTCCTGGAGCTTTAC
AB227	Forward. Check *TTC18XX* mutation	GAAGCCCCGGTTAGGGAG
AB228	Reverse. Check *TTC18XX* mutation	GAGAGAGCCTCGGAGCGATG
AB229	Forward. Construction of pAB214, pAB221. Fusion of sYFP and sGFP reporter to C-terminal of TtAgo	ccatggGTCCAGGCCGTGCTGGTC
AB230	Reverse. Construction of pAB226, pAB227, pAB214 and pAB221. Fusion of sYFP and sGFP reporter to C-terminal of TtAgo	actagtAACGAAGAAGAGCTTTTCCCGGTC
AB231	Reverse. Construction of pAB175, pAB176. Fusion of sYFP and sGFP reporter to TtHerA	actagtCCTGAAGAACTCCCGGCGCAG
AB232	Forward. Construction of pAB175, pAB176. Fusion of sYFP and sGFP reporter to TtHerA	CCATGGGTGAAGCGTATCGGCGTG
AB233	Forward. Construction of pAB183 and pAB184	aaactgcagCTTGGAGGAAGCCCGGAAG
AB234	Reverse. Construction of pAB183 and pAB184	aaagaattcGAGGGAGGAGTAGCGCCATC
AB235	Reverse. Construction of pAB187 and pAB188	atcaagcttGAGGTGTAGGGCCTAAAGACCAC
AB236	Forward. Construction of pAB187 and pAB188	aaagaattcCATAGCCGTGCCCCGTTTC
AB237	Forward. Construction of pAB185 and pAB186	aaagaattcCTCATTGAGCTGGCGATCGTC
AB238	Reverse. Construction of pAB185 and pAB186	atcaagcttGTATCGGCGCTCGCGTTG
AB239	Reverse. Construction of pAB181 and pAB182	atcaagcttGAGGCCTTTGACCTGGACCTC
AB240	Forward. Construction of pAB181 and pAB182	aaagaattcGAGGTGAGGATGTCCACGCTG
AB241	Forward. Sequencing 16S (0-30F)	GAGTTTGATCCTGGCTCAG
AB242	Reverse. Sequencing 16S (530R*)	GTATTACCGCGGCTGCTG
AB243	Reverse. Check religated *T. thermophilus* clones from *kat* cassette after transposition by reverse PCR	GTGCTCGTACCCCTCCTCCTC
AB244	Forward. Check religated *T. thermophilus* clones from *kat* cassette after transposition by reverse PCR	GTACGGGGCAAGAAGGACCTTC

AB245	Reverse. Check religated *T. thermophilus* clones from *kat* cassette after transposition by reverse PCR	CACCTTCCACTCACCGGTTGTC
AB246	Forward. Check religated *T. thermophilus* clones from *kat* cassette after transposition by reverse PCR	GATTCAGGAGTGGACAGAACGACAC
247	Forward. Construction of pAB201	aaacatatgACGGACCAGGAACTTCTGCA
248	Reverse. Construction of pAB201	aaaaagcttTCATGTCCGCACACGCCCT
249	Forward. Construction of pAB207 and 208. Upstream amplification of *ttherA*	aaagaattcGTAATCTAGGGGCATGGCCTG
250	Reverse. Construction of pAB207 and 208. Upstream amplification of *ttherA*	aaatctagaCTTCACCTCGCACCTCCCA
251	Forward. Construction of pAB207 and 208. Downstream amplification of *ttherA*	aaatctagaGATGGCCTCGCCATGAAGA
252	Reverse. Construction of pAB207 and 208. Downstream amplification of *ttherA*	AAACTGCAGCAGGGGCGAATAC
253	Forward. Construction of pAB210 and pAB213. Fusion of sYFP and sGFP reporter to C-terminal part of TtCptA	aaaccatggGAAGACCGGGTGCACTACCTG
254	Reverse. Construction of pAB210 and pAB213. Fusion of sYFP and sGFP reporter to C-terminal part of TtCptA	aaaactagtTGTCCGCACACGCCCTTG
255	Forward. Construction of pAB224 and pAB225. Fusion of sYFP and sGFP reporter to TtCptA	aaaccatggATGACGGACCAGGAACTTCTGCA

ANNEX III

TABLE III. Buffers and solutions employed in this dissertation

Name	Composition	Use
A1 solution	K_2HPO_4 430 mM, KH_2PO_4 185 mM, $(NH_4)_2SO_4$ 1890 mM	Preparation of incubation plates for twiching-motility assays
Adhesion washing solution	Tris-HCl 25 mM, NaCl 100 mM	Cell wash in sub-agar-surface adhesion tests
Anode buffer I	Tris 300mM pH 10.4, methanol 10 %	Acrilamide gradient gel electrophoresis and immunoblotting with nitrocellulose membranes
Anode buffer II	Tris 25 mM pH 10.4, methanol 10 %	Acrilamide gradient gel electrophoresis and immunoblotting with nitrocellulose membranes
ATPase activity buffer	Tris 100 mM pH 7.5, EDTA 4 mM	Determination of hydrolized ATP by biological material in ATPase activity assays
B solution	$MgCl_2$ 60 mM, $CaCl_2$ 17 mM	Preparation of incubation plates for twiching-motility assays
C solution	$FeSO_4$ 216 mM, $CoCl_2$ 34 mM, $NiCl_2$ 0.85 mM in 0.01 N H_2SO_4	Preparation of incubation plates for twiching- motility assays
Cathode buffer	Tris 25 mM pH 9.4, methanol 20 %, $C_6H_{13}NO_2$ 40 mM	Acrilamide gradient gel electrophoresis and immunoblotting with nitrocellulose membranes
Chemiluminiscent development solution	Tris-HCl 100 mM pH 8, luminol 1.25 mM, luciferine 36.5 µM and H_2O_2 0.022 % (p/v)	Western-blot membrane development
Coomassie dye	Methanol 45 %, acetic acid 10 % (v/v), Coomassie Brilliant Blue-R 0.25 % (w/v)	Protein staining for SDS-PAGE
Crystal violet	Crystal violet 0.1 %	Staining cells in biofilm formation assays
Destaining solution	Methanol 45 %, acetic acid 10 % (v/v)	Protein de-staining for SDS-PAGE
DNA 10X loading buffer	TAE 10X, glycerol 30 % (v/v), bromophenol blue 0.25 % (w/v), xylene cyanol FF 0.25 % (p/v) .	DNA sample preparation for agarose electrophoresis
DNase I buffer	Tris-HCl 100 mM pH 7.5, $MgCl_2$ 25 mM, $CuCl_2$ 1mM	Prevent DNA uptake in conjugation assays. Extracellular vesicle purification
Elution buffer ATPase	Tris-HCl 25 mM pH 7.5; filtered and sonicated	Buffer dilution of TtHerA and TtCptA for ATP incubation and further EM analysis
F solution	$MnCl_2$ 25.3 mM in 0.01 N HCl	Preparation of incubation plates for twiching and motility assays
IB elution buffer	Phosphate buffer 25 mM pH 7.5; 200 mM imidazole; 1 mM $MgCl_2$	Inclusion bodies protein purification elution buffer
IB equilibration buffer	Phosphate buffer 25 mM pH 8; 6 M urea	Inclusion bodies protein purification equilibration buffer
IB resuspension buffer	Phosphate buffer 25 mM pH 8; triton X-100 0.2 % (v/v)	Inclusion bodies protein purification resuspension buffer

IB washing buffers	Phosphate buffer 25 mM pH 7.5; 4-0 M urea; 1 mM $MgCl_2$	Inclusion bodies protein purification gradient washing buffer
Laemmli 5X lysis buffer	Tris-HCl 300 mM pH 6.8, SDS 5% (v/v), β-mercaptoethanol 10 %, glycerol 50% (v/v), bromophenol blue 0.002%, EDTA 25mM.	Protein sample preparation for SDS-PAGE
Miniprep solution I	Tris 25 mM pH 7.5, Glucose 50 mM, EDTA 10 mM	Alkaline lysis plasmid DNA purification
Miniprep solution II	NaOH 0.2 M, SDS 1 % (v/v)	Alkaline lysis plasmid DNA purification
Miniprep solution III	Potassium acetate 3 M, glacial acetic acid 0.5 M	Alkaline lysis plasmid DNA purification
PBS 1X buffer	NaCl 150 mM, KCl 25 mM, Na_2HPO_4 8mM, KH_2PO_4 1.5 mM	Cell extract resuspension and wash for microscopy preparations
Ponceau S	Ponceau Red 0.2 % (w/v), acetic acid 30 % (v/v)	Red stain for immunoblot detection using nitrocellulose membranes
Rnase A buffer	Tris- HCl 10 mM pH 7.4, NaCl 0.1 M, glycerol 20 % (v/v)	RNA removal from samples
Running buffer 5X	Tris 125 mM pH 8.8, glycine 1 M, SDS 20 mM	Denaturant protein electrophoresis
SOB medium	Tryptone 2% (w/v), yeast extract 0.5 % (w/v), NaCl 10 mM, KCl 2.5 mM, $MgCl_2$ 10mM, $MgSO_4$ 10mM, MqH_2O to 1 L	Broth medium for *E. coli* competent cells
SOC medium	Tryptone 2% (w/v), yeast extract 0.5 % (w/v), NaCl 10 mM, KCl 2.5 mM, $MgCl_2$ 10mM, $MgSO_4$ 10 mM, glucose 20 mM, MqH_2O to 1 L	Catabolite repression medium. Transformation of *E. coli* cells
Stripping buffer	Tris-HCl 62.5 mM pH 6.7, β-mercaptoethanol 100 mM, SDS 2% (v/v)	Removal of tightly bound antibodies from Immunoblot membranes
TAE 1X	Tris-acetate 40 mM pH 8; EDTA 1 mM	Agarose electrophoresis
Talon elution buffer	Phosphate buffer 50 mM pH 7, NaCl 300 mM, imidazole 150 mM	Elution buffer in His-tag purification protocols using TALON Co^{2+} resin
Talon Equilibration buffer	Phosphate buffer 50 mM pH 8, NaCl 300 mM	Equilibration buffer in His-tag purification protocols using TALON Co^{2+} resin
Talon washing buffer	Phosphate buffer 50 mM pH 7, NaCl 300 mM, imidazole 10 mM	Washing buffer in His-tag purification protocols using TALON Co^{2+} resin
TB medium	Tryptone 0.8 % (w/v), yeast extract 0.4 % (w/v), NaCl 50 mM, *Thermus* H_2O to 1 L, pH 7.5	*Thermus* growth medium
TBS-T	Tris-HCl 20 mM pH 8, NaCl 150 mM, Tween-20 0.1% (v/v).	Blocking, washing and antibody diluting solution in Western-blot assays
TE	Tris 10 mM pH 8, EDTA 1 mM	Cell extract wash
TES	Tris 50 mM pH 7.5, EDTA 10 mM, sucrose 25 % (w/v)	Cell wash during chromosomal DNA extraction and purification
TFBI	RbCl (100 mM), $MnCl_2 \cdot 4H_2O$ (50 mM), K Acetate (30 mM), $CaCl_2 \cdot 2H_2O$ (10 mM), glycerol (15%) pH 5.8	Competent cells preparation buffer
TFBII	MOPS (10 mM), RbCl (10 mM), $CaCl_2 \cdot 2H_2O$ (75 mM), glycerol (15 %) pH 7	Competent cells preparation buffer
Transfer buffer	Tris 48 mM pH 8.5, glycine 39 mM, SDS 0.037% (w/v) and methanol 20 %	Protein transfer to semi-dry PVDF membrane

TS 1X	Tris-HCl 50 mM pH 7.5,NaCl 50 mM	Cell extract wash and resuspension
Wash solution	Tris-HCl 50 mM pH 7.5, EDTA 5 mM, NaCl 50 mM	Cell wash during chromosomal DNA extraction and purification

ANNEX IV

SUPPLEMENTARY MATERIAL FOR CHAPTER 4 "HORIZONTAL TRANSFER OF VESICLE-PROTECTED eDNA AMONG *Thermus* spp"

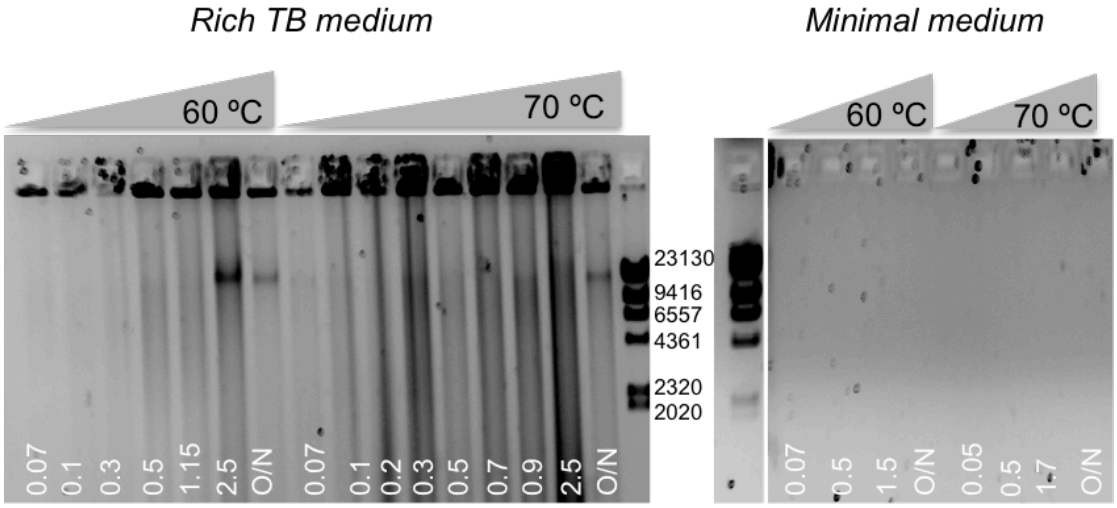
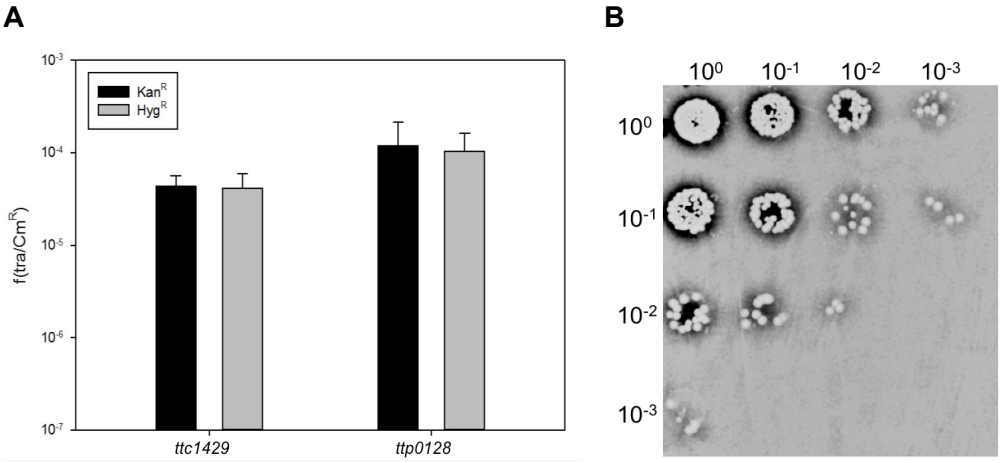
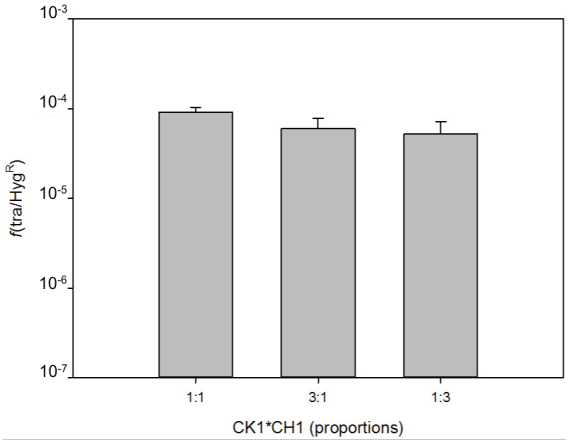


FIGURE S4.1. eDNA production is related to growth rates in rich TB medium but not in mineral medium. eDNA samples were prepared from similar cell volumes at different optical densities from cultures of the HB27wt growing at 60 ºC and 70 ºC in TB rich medium. In parallel, HB27wt was grown in mineral M162 medium. 10 µl of the precipitated eDNA were run in 1 % agarose gels and stained with ethidium bromide, Molecular weight sizes are expressed in bp.

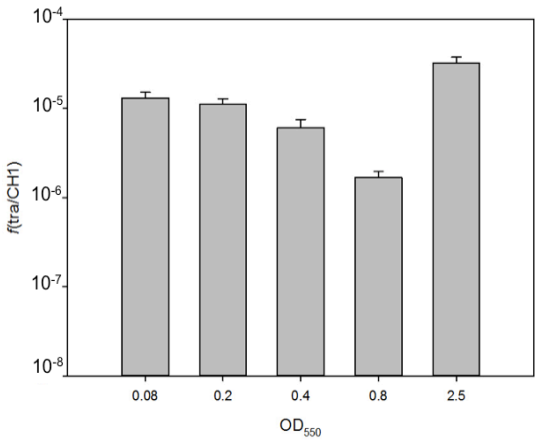
SUPPLEMENTARY MATERIAL FOR CHAPTER 5 "CONJUGATION IN *Thermus thermophilus*"



C



D



E

CH1	10	20	30	40	60 (min)
CK5	++	++	++	++	++
CK6	++	++	++	++	++
CK1	++	++	++	++	++
CK7	++	++	++	++	++
CK8	++	++	++	++	++

F

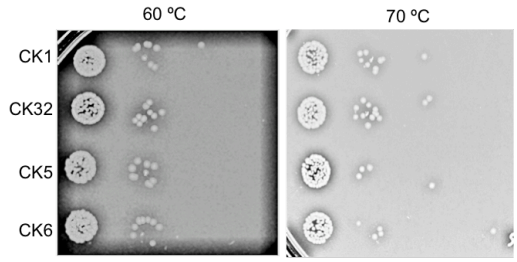


FIGURE S5.1. Conjugation is not dependent on the selection marker, the dilution employed, the incubation time of the mixing nor the temperature set for the experiment. Mating experiments involving several HB27 derivatives. **A.** Isogenic CK32 and CH14 strains, harbouring a *kan* and an *hyg* resistance at loci *TTC1429* and isogenic strains PK31 and PH21, at loci *ttp0128*, were mated with Cm^R strain 27CM. Bars represent the transfer frequencies of the Kan^R (black) and the Hyg^R (grey) of these matings (*p*-value: 0.765 and 0.643, n=3). **B**. 10-µl sample from several dilutions of matings between CK1*CH1 were plated on TBHK. **C.** Transfer frequencies obtained from matings involving CK1and CH1 strains at different cell mass proportions (X axis) (*p*-value: 0.06; n=3). **D.** CK1 and CH1 were conjugated along its growth cycle, referred as optical densities (OD_{550}, X axis) and transfer frequencies represented by bars (*p*-value< 0.01; n=3). **E**. Schematic representation of the transconjugants grown from several matings involving CH1 and different Kan^R strains (CK5,CK6, CK1, CK7, CK8) incubated together in different times (referred in minutes). **F**. Transconjugants grown in TBHK plates at 60 ºC (left-side panel) and 70 ºC (right-side panel) of different mixes of saturated cultures of CH1 with different Kan^R strains CK1, CK32, CK5 and CK6.

SUPPLEMENTARY DATA FOR CHAPTER 6 "EXPLORATION OF DNA PUSHING PROCESS DURING *T. thermophilus* CONJUGATION"

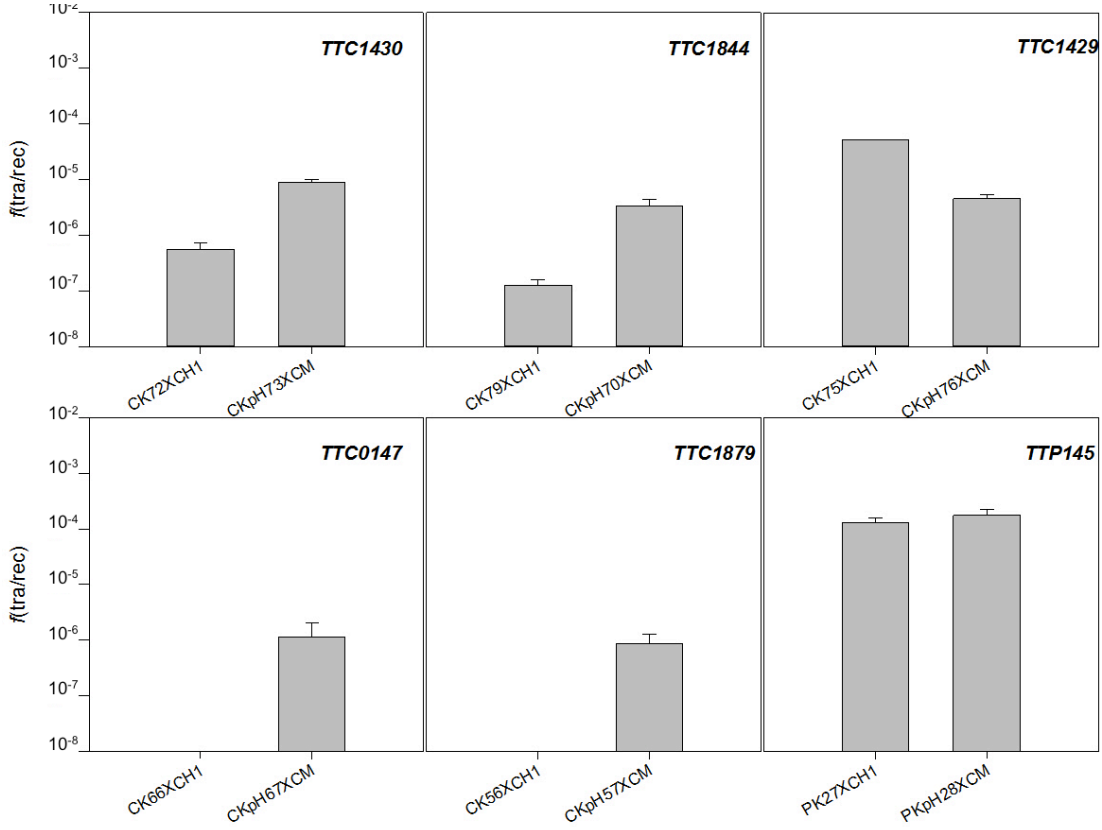


FIGURE S6.1. Complementation assays of the candidate proteins to be involved in DNA pushing. Conjugation frequency rates referred to the wild type strain of matings involving Δ*push::kat*Δ*pilA4* Kan^R derivatives and its complementation derivative (Δ*push::kat*Δ*pilA4*^{s}*push::hyg)* from each of the 6 candidates crossed with wild type control strains (CH1 and Cm^R (CM), respectively) (n=3).

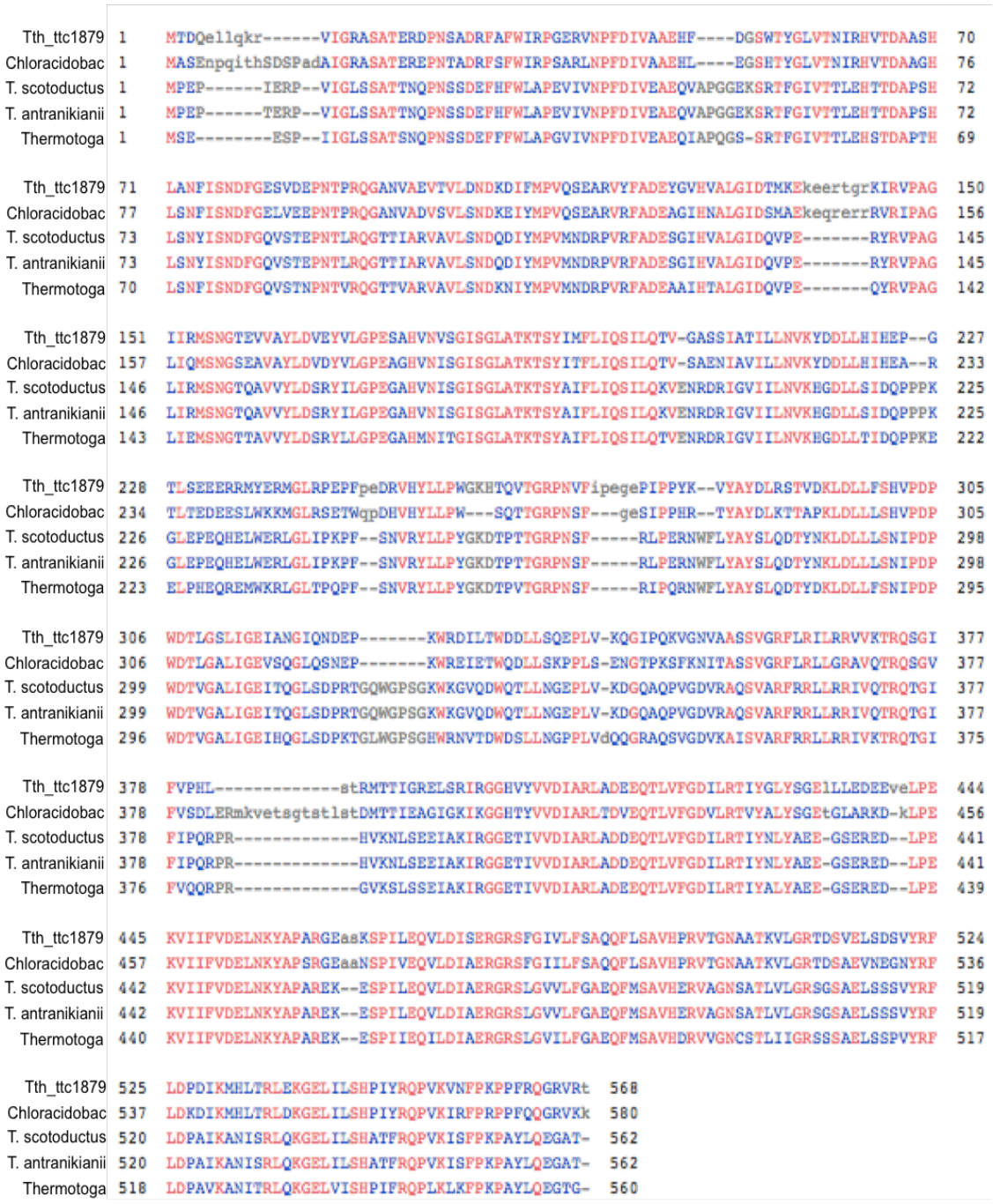


FIGURE S6.2. Multiple sequence alignments of CptA *(TTC1879).* P-BLAST results for CptA with its best homologues. These homologues were ordered by highest similarity to CptA and aligned with COBALT: Chloracidobac (*Chloracidobacterium thermophilum*), T. scotoductus *(Thermus scotoductus SA01),* T. antranikianii (*Thermus antranikianii*) and *Thermotoga (Thermotoga napholitana).* Common aminoacids to the 5 sequences are represented in red.

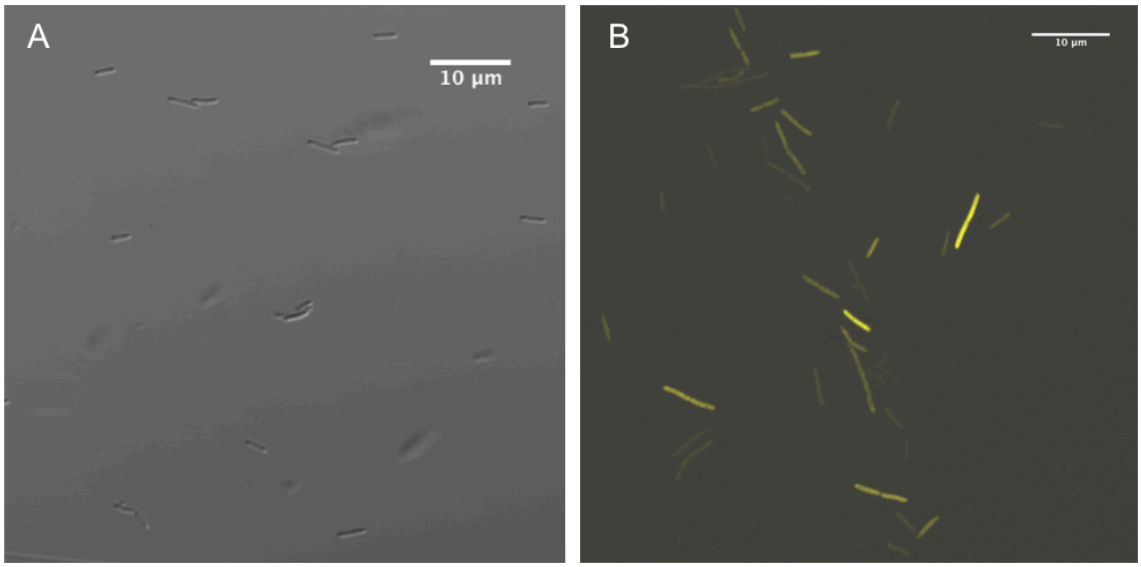


FIGURE S6.3. Fluorescent proteins distribute homogeneously in *T. thermophilus*. Overlapping images of bright field and yellow channel from HB27^{EC} preparations (A) and PpH27, harbouring the plasmid pMHPnqosYFP (B), working as negative and positive controls, respectively.

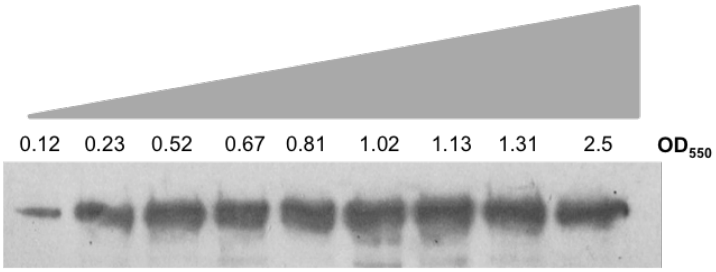


FIGURE S6.4. Expression of CptA in *T. thermophilus* HB27. Detection of CptA expression along its growth cycle. The derivative CH29, harbouring a sYFP reporter fused to CptA, was grown at 60 ºC and several samples were taken along the growth (OD_{550}) and and analysed for the presence of CptA by Western-blot. Primary antibody was monoclonal α-GFP (1:5000). All samples contained equal cell mass.

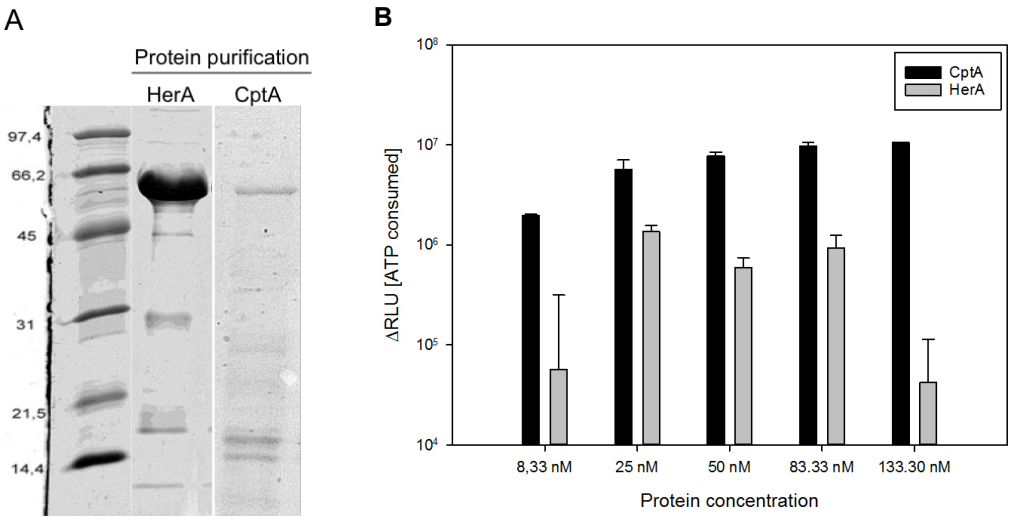
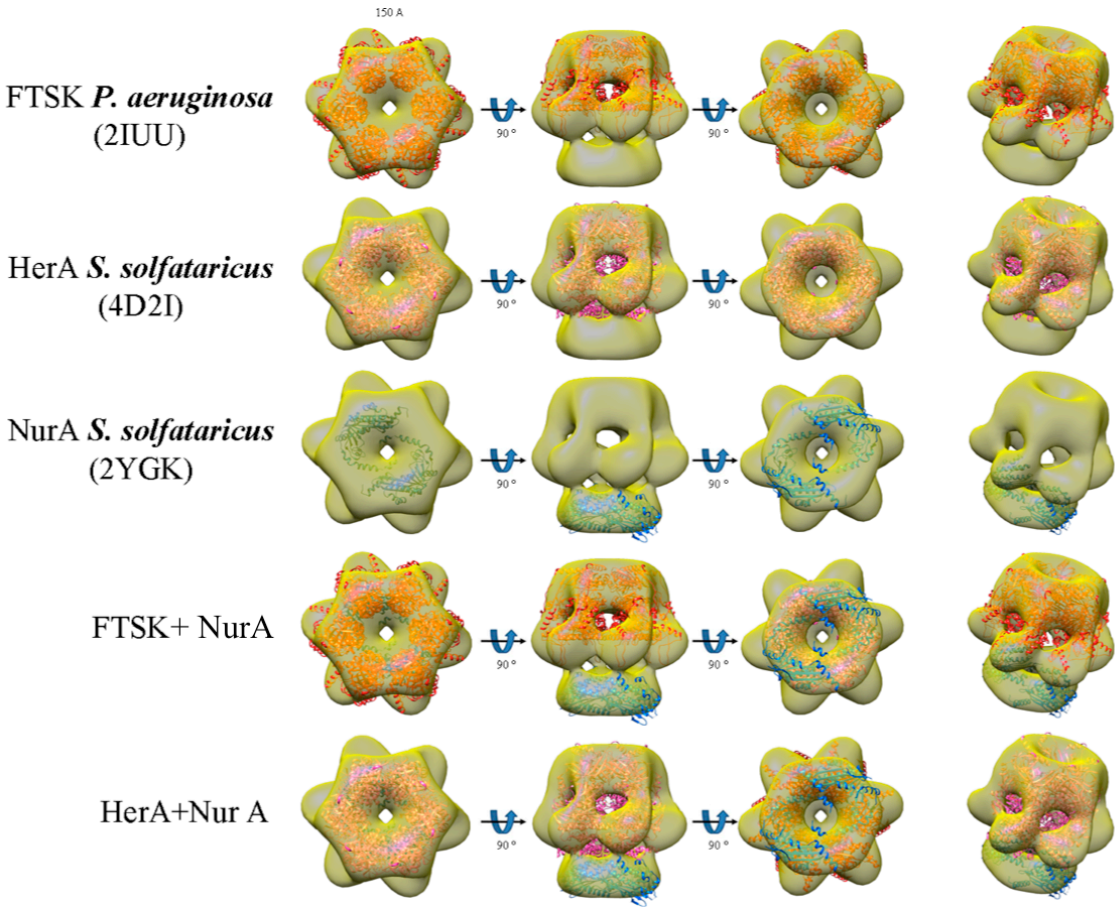


FIGURE S6.5. Protein purification and ATPase activity assays. A. SDS-PAGE of HerA and CptA eluted fractions after purification by NI-NTA affinity chromatography. **B.** ATPase activity of purified proteins CptA and HerA (26 nM) where ATP consumed is expressed in luminiscense emitted as RLU/s per concentration of ATP employed, using CLSII-Bioluminiscent ATPAssay kit (Roche).

A



B

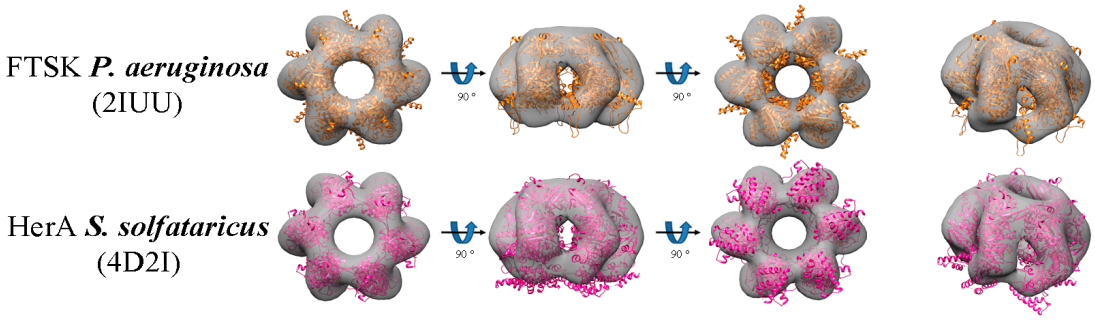


FIGURE S6.6. Fitting of HerA and CptA with different proteins. 3D models calculated for proteins HerA and CptA hexamers were juxtaposed with available structures at the PDB database harbouring equivalent motifs. **A.** HerA 3D model was used to fit FtsK protein structure from *P. aeruginosa*, and HerA and NurA from *Sulfolobus solfataricus.* Combinations of all three templates were included too in the fitting. **B**. CptA was fitted with *P. aeruginosa* Ftsk and HerA from *S. solfataricus.*

TABLE S6.1. Mutagenesis screening of conjugation in *T. thermophilus* HB27. Screened proteins as potential candidates selected for individual analysis prior to mutagenesis screening for proteins involved in conjugation in *T. thermophilus*. Grouped by their pFAMs, candidates encoded in the chromosome (A) or pTT27 megaplasmid (B) were listed. Note that some loci are clustered more than once, classified in several pFAMs.

A

PFAM	Loci (chromosome)
AAA_23 (PF13476)	TTC0543, TTC0922, TTC1161
UvrD_C_2 (PF13538)	TTC0638, TTC1062
AAA_25 (PF13481)	TTC0173, TTC0363, TTC0700, TTC1075, TTC1106, TTC1890
AAA_10 (PF12846)	TTC0147, TTC0174, TTC0474, TTC1430, TTC1879
AAA_31 (PF13614)	TTC0295, TTC0879, TTC1605
AAA_21 (PF13304)	TTC0071, TTC0420, TTC0543, TTC0573, TTC0648, TTC0795, TTC0843, TTC0928, TTC1043, TTC1071, TTC1095, TTC1527, TTC1559, TTC1723, TTC1890, TTC1934
T2SE (PF00437)	TTC1415, TTC1622, TTC1844
T2SE_G (PF08334.7)	TTC1836, TTC1839
UvrD_C (PF13361)	TTC0638, TTC1062
UvrD-helicase (PF00580)	TTC0638, TTC1062
AAA_2 (PF07724)	TTC0174, TTC0251, TTC0264, TTC0746, TTC1123
AAA (PF00004)	TTC0174, TTC1123, TTC1135, TTC1523
DEAD (PF00270)	TTC0533, TTC0893, TTC0902, TTC1895
Helicase_C (PF00271)	TTC0533, TTC0893, TTC1531, TTC1895
PrimPol (PF09250.7)	TTC0656
NurA (PF09376.6)	TTC1429

B

PFAM	Loci (pTT27)
Helicase_C_2 (PF13307)	TTP0128
AAA_14 (PF13173)	TTP0140, TTP0211
UvrD_C_2 (PF13538)	TTP0191
DUF4143 (PF13635.1)	TTP0081
RPA (PF10134.4)	TTP0085, TTP0145
AAA_16 (PF13191.2)	TTP0084
XerD (PF02899.13)	TTP0208

SUPPLEMENTARY DATA FOR CHAPTER 7 "PUTATIVE ROLE OF INSERTION SEQUENCES IN CONJUGATION AND DEVELOPMENT OF A TRANSPOSON-BASED MUTAGENESIS SYSTEM IN *Thermus thermophilus*"

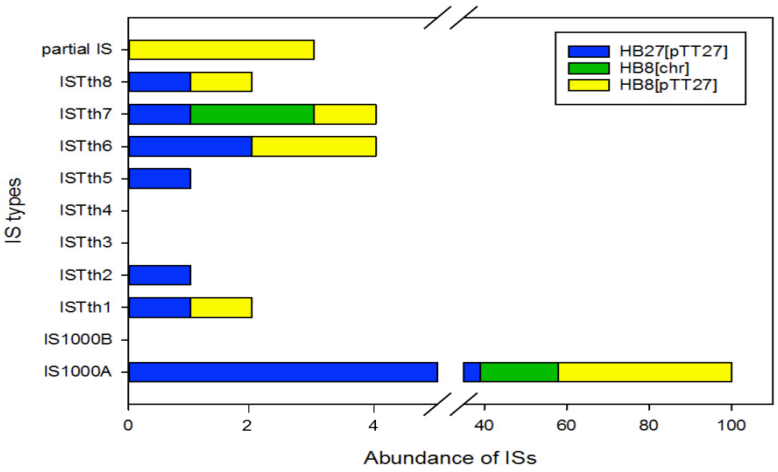


FIGURE S7.1. Abundance of partial ISs among *T. thermophilus* genomes. The number of copies of partial ISs found in either HB27 and HB8 genomes was classified upon their specific compartment (pTT27 and chromosome) and by defined typologies.

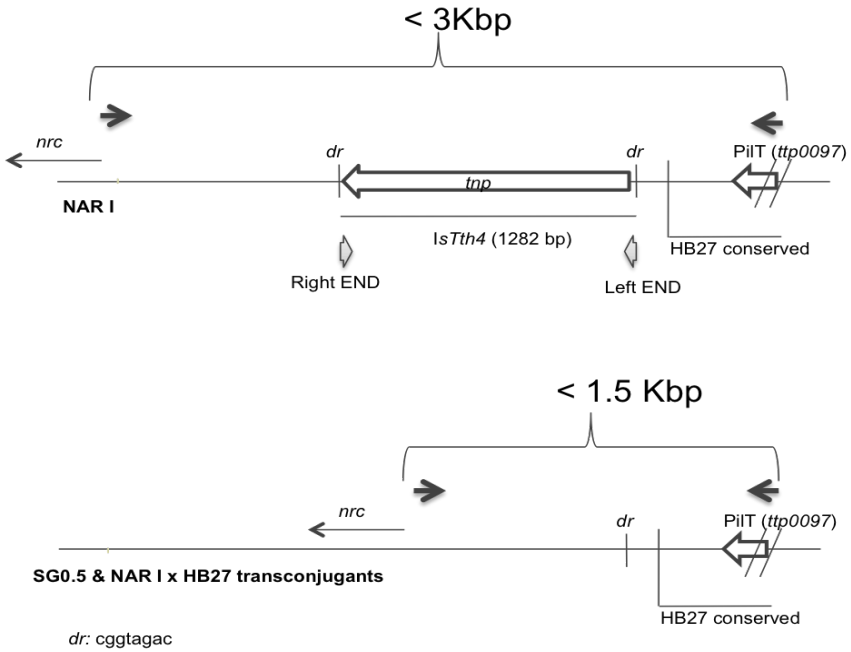


FIGURE S7.2. Schematic representation of the PCR based experiment. The neighbouring region of the NCE element is represented, indicating with black arrows the annealing site of PCR primers For, located in *TTP097* gene, and the two Rev primers. Conserved regions in HB27 as well as the presence of an *ISTth4* its subsequent direct repeats (dr) and inverted repeats (grey arrows) are also portrayed.

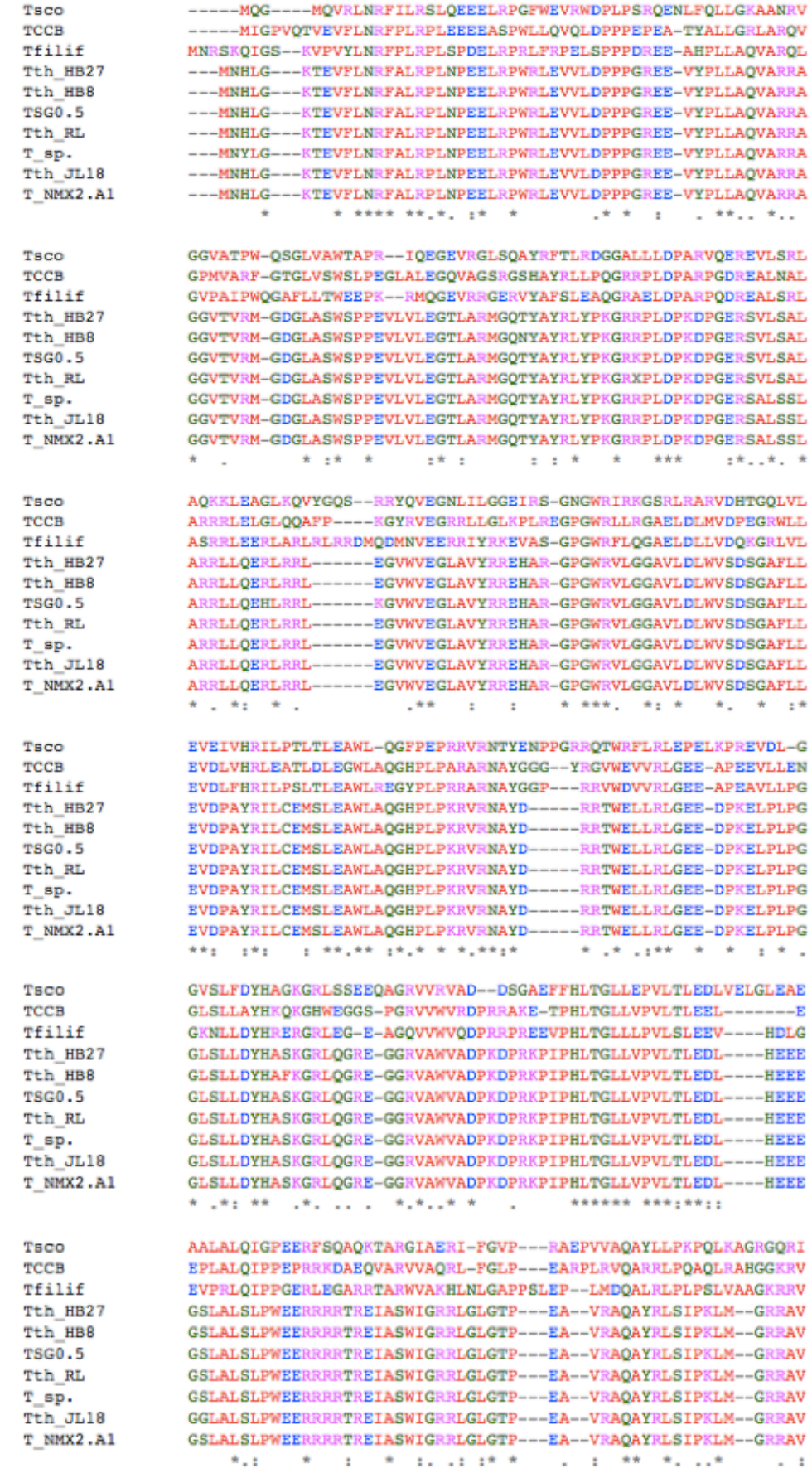
TABLE S7.1. Main characteristics of *ISBstI* and *ISNth2* employed for the development of the mutagenesis system.

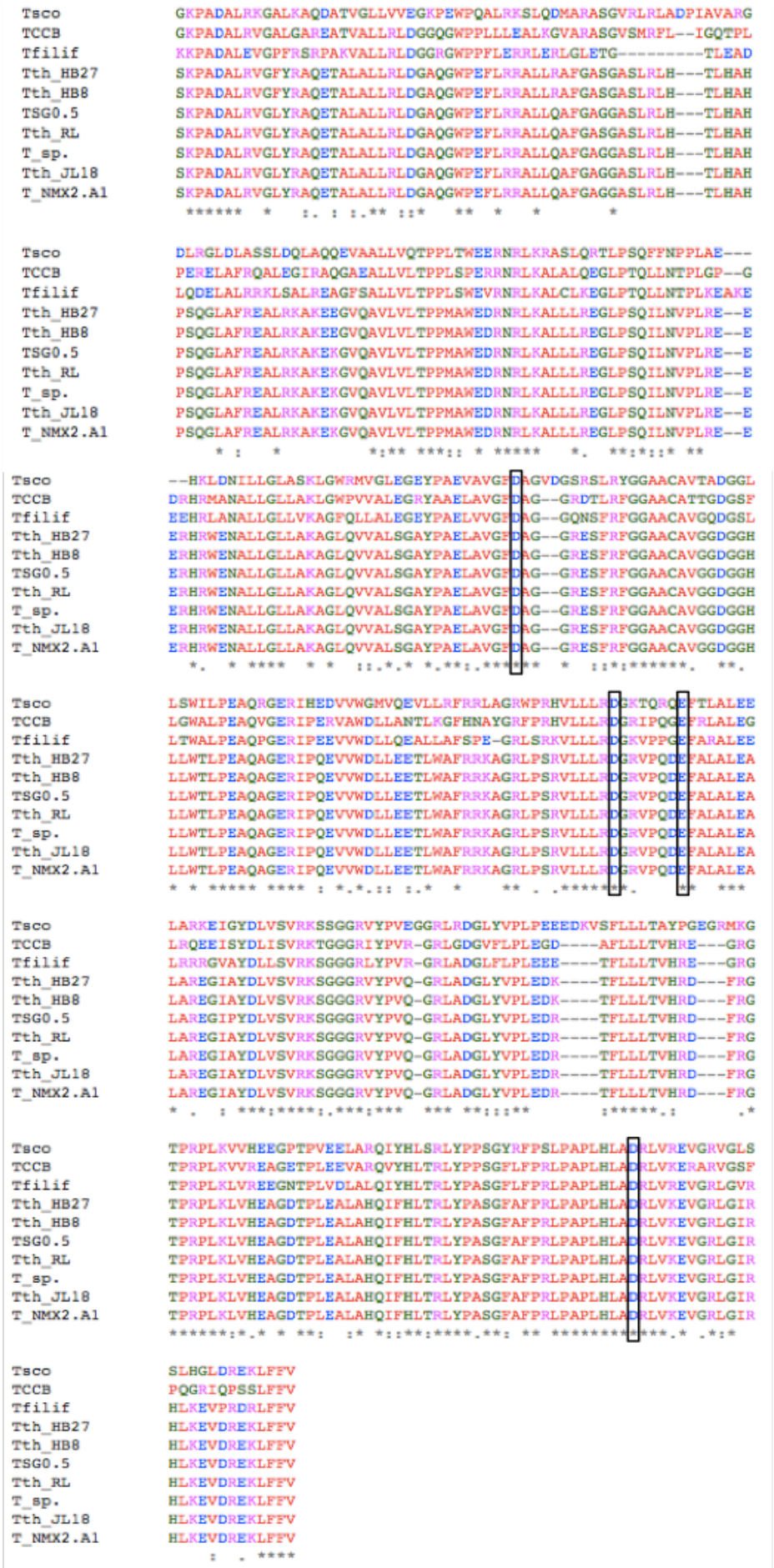
Name	ISNth2	ISBstI
Organism	*Natranaerobius thermophilus*	*Bacillus stearothermophilus*
Taxonomy	Bacteria, Firmicutes, Clostridia, Natranaerobiales, Natranaerobiaceae, Natranaerobius	Bacteria, Firmicutes, Bacillales, Bacillaceae,
Growth Tº	53 ºC	65 ºC
Family	ISLre2	IS481
Length	1561	1461
IR length	15/22	7
DR length	9	-
ORF tranposase length	462(129-1517)	417(165-1418)
IR identities	93%	88%
Left end	GTTATTGTAAAATTACTGTAGGATTTTTT GGAGAAAGCAGGAAGGAAAGG	TGTTTCTCCAACGTCCGGATCCTGATTT TACATACATATCCTAAATTGGA
Right end	GATCGTGTCAAGCCACTGTAGGTTAAAA AGAAACCTTAATTCTCCTAGAT	TGTTTCTTTCGCAAACGGCTCCCGCGA AAGGGAGTAAAACGATTTATACA
INSERTION sites	GTCTATTAAG (AAAAGTATT) ACTTCGTTTGA//ATTTCAACAA (AACTTTTAG) ATTTCCTTTGA//AAAAACCTAA (AAAGCTTAC) ATATGGCTTGA	-
IS_seq	GTTATTGTAAAATTACTGTAGGATTTTTT GGAGAAAGCAGGAAGGAAAGGAAGCAA AAAAGAAGAGGAACCTCACATTCTCCGC ACGAGAATGAAGTATACCCAAAAAAGTA AGGAGTGAGGTTCCTCATGGAAATAATT CAACTTGTAGAAGAAAAAATCCTTCAAGT TTCAGAAAAAGTGTTAAATGTTTTAGACG GAGAAATAGAGTACGATACATTCACAAA GCTACTAAAAGAAGAGCTAGACGGTTTA GGTAGTGATATTCTTAAAGAAGTACTAG AAGCATTGGATGCCCAAATAAAAGAAAA TAGAAGAGATCGTAAAGGTTGGGTCGTA GAAAGGAAAGCTGACCCCAAGAGTATCT TAACCCCATTTGGTGATATGGTATACAA CCGAACATATTATAAAAATAAAGAAACAG GTGAATATAAATATCTAGCTGATGAAAAA GCTGGCATAACCTCACATATGCGAGTAG ATTCTACTTTAAAATCAGATATAGCTGAT GCAGCCACTAAGGTGTCGTATGAGAAAG CGACAGAAGAAGTGAGTAGGTTTAACCA AGACCTAAAATTAAGTAAACAGACAGTA GCCAATACAGTGAAAGAATTTGAATTAG AACCTCTAGAGCAACCAGAAGAAAAGAA AAAAGTATCAAACCTTTACATAGAGGCA GATGAAGATCATTTGTCTATGCAAACTTC GAAGAGATCTGAAACAAAGCTAATTTAT GTCCATGAAGGAATAAAAGGTAAAAAAC GTAGATCTCTTAAAAATTGTCAACATTTC ACAACAATAACTGAATCCCCGGATAAAT TTTGGTTAACAGTATGTGACTATATCGAA TCACATTATGACACAAAGGACACAAATAT TTTCATATCAGGAGATGGGGCTAAATGG ATCCGAGTAGGTGAAGAATATATCCCAA ATGCGACATATATATTGGATAAATTTCAT CTATCCAAATACATAATAGCAGCTACCG GTCATGCCCCAAAGCTAAGGAGACAGAT ATATAAGAACATTAAAAACCTAGATCAAG AAGCAGTATTTGAAAATTTACAAGAAGCA	TGTTTCTCCAACGTCCGGATCCTGATTT TACATACATATCCTAAATTGGAATGGTT GAGGTTCCCACAAACCTTTCCTATCGA CGGAGAAAAAGCGATCCGGTACGATAA GGGCAGACGGACCGGAACGGTCCTGA TTCCAAGTGAGACAAGGAGGAGATTCG AATGGATGAGTCGATGAGACACGACAT TGCCCTGTTTCGGTACGGGCTAATCGC TCCGTTGGTGAATGGGCAGGTGGAACC GAAGGCGTATCTGAAAGAGGTGAGCGA GCGAGTGCACCATGTTCCCCACCAAGG CGACAAACGAATCGCAGCCAAGACGAT CCTCGATTGGTGCACCCGATACAAAAA AGGGGGCTTTGACGCCTTAAAGCCGAA GCGCCGTTCGGACCGCGGCCACTCAA GGCGTCTGTCGCCCGATGATGAGGATC ATATTTTAGCCCTAAGAAAAGAACATCC CACCATGCCCGTGACGGTGTTCTACGA ACACCTGATCGAGCAGGGGGAAATCCC AAAAAACCATATCTCTTACTTTACTATAT ACCGATTGTTGAAAAAACACAACCTTGT CGGGAAAGAAATCTTGCCCATGCCGGA GCGAAAGCGCTTTGCGTATGACCAGAT CAATGAGCTTTGGCAAGGGGACTTATC CCATGGACCAACGATTCGAGTCAATGG GAAAGCCCAGAAAACGTTTTTGATCGC CTATATCGATGACTGCTCGCGGTTTGT GCCCGTAACGTATAGTTTCCCTTCGGA GAAGTTTGACGGGCTGCGGATCGTCAC GAAGGAAGCGGTGCTTCGTTGCGGGA AGCCGAAGCGCATTTACTCGGACAACG GGAAAATTTATCGATCCGAGGTGCTGC AGTATGCGTGCGCCGAGATGGGAATTA CGCTCATCCACACCCAGCCGTATGACC CGCAAAGCAAAGGGAAAATCGAACGGT TCTTCCGCACCGTACAGACGCGGTTTT ATCCGCTGATGGAGTTGGATCCCCCGA AGTCGCTCGAATACCTGAACGAGCGTT

IS_pep	CTTACAAAGGCAGATACAGAAGCTAGAC AGAAACGCATAATGGCAACAATTAAATAT ATTAAAAACAACTGGGATGGCATAGAGG CATCAGTAAACCACCCTGAAGTAGGTTG TAGTGCTGAAGGCCATGTAAGCCATATT TTAGCATCAAGAATGAGTAGTCGACCAA TGGCCTGGAGTGAAAAAGGCGCAACTA AAATGGCCGAGATGTTGGCTACAAAAGC TAATGGAAAATCAGTAAAAGAAGCTTACT TATCCACAAAAGGTCGTCAAGAAGCTGA AATAGTTAACCTACAGAACCAAATAAAAC AAGAACTAAAAAAACTTACAACTAAAAAG AAACTAGGAAAAGCCAACAATGGCAATG TACCTTTGATGAATGGAAAGTACAACTTA ACAAGAACGGCTATAAAAGGGTTGAACA GGAAACAGATTATCTAGGAGAATTAAGG TTTCTTTTTAACCTACAGTGGCTTGACAC GATC MEIIQLVEEKILQVSEKVLNVLDGEIEYDTF TKLLKEELDGLGSDILKEVLEALDAQIKENR RDRKGWVVERKADPKSILTPFGDMVYNR TYYKNKETGEYKYLADEKAGITSHMRVDS TLKSDIADAATKVSYEKATEEVSRFNQDLK LSKQTVANTVKEFELEPLEQPEEKKKVSN LYIEADEDHLSMQTSKRSETKLIYVHEGIK GKKRRSLKNCQHFTTITESPDKFWLTVCD YIESHYDTKDTNIFISGDGAKWIRVGEEYIP NATYILDKFHLSKYIIAATGHAPKLRRQIYK NIKNLDQEAVFENLQEALTKADTEARQKRI MATIKYIKNNWDGIEASVNHPEVGCSAEG HVSHILASRMSSRPMAWSEKGATKMAEM LATKANGKSVKEAYLSTKGRQEAEIVNLQ NQIKQELKKLTTKKKLGKANNGNVPLMNG KYNLTRTAIKGLNRKQII	TTTGGAGGTGGCTTGAAGAGGAGTATC ATCGAAAACCGCACGCCTCGCTGGACG GAAAAACGCCGCATGAGGTGTTTCAAT CGCAAGTGCATCTCGTGTCGTTCATCG AAGATGGCGATTGGCTGGATGCGATTT TTCTGAAGCGGGAGCAACGCAAGGTCA AAGCCGATGGAACGATCACCTTGAACA AGCAGCTGTATGAAGTGCCGCCCCGGT TCATTGGGCAATCGATTGAACTCCGCT ATGACGAACGGGGCGTCTATGTGTACG AAGATGGAAGGAAGGTGGCGGAGGCC GTTTGGGTTCGTTTGGAGGACAACGCC TATGTAAAACGCCATCGGTCCCCGTTT GCGGCGATTCCGGCGAAGGAGGGAGA ACACGGTGTATAAATCGTTTTACTCCCT TTCGCGGGAGCCGTTTGCGAAAGAAAC A MDESMRHDIALFRYGLIAPLVNGQVEPKA YLKEVSERVHHVPHQGDKRIAAKTILDWC TRYKKGGFDALKPKRRSDRGHSRRLSPD DEDHILALRKEHPTMPVTVFYEHLIEQGEI PKNHISYFTIYRLLKKHNLVGKEILPMPER KRFAYDQINELWQGDLSHGPTIRVNGKA QKTFLIAYIDDCSRFVPVTYSFPSEKFDGL RIVTKEAVLRCGKPKRIYSDNGKIYRSEVL QYACAEMGITLIHTQPYDPQSKGKIERFF RTVQTRFYPLMELDPPKSLEYLNERFWR WLEEEYHRKPHASLDGKTPHEVFQSQV HLVSFIEDGDWLDAIFLKREQRKVKADGT ITLNKQLYEVPPRFIGQSIELRYDERGVYV YEDGRKVAEAVWVRLEDNAYVKRHRSP FAAIPAKEGEHGV
% G+C	35.6	52.2
nº complete copies	3	17
Protein accesion number	WP_012447610.1	ABI49958.1
Comment	Species-specific. N. thermophilus JW/NM-WN-LF	Multiple copies of both complete and partial IS are found in several *Geobacillus* spp and in plasmid pGS18 in
Homologues	HP* in *T. xylanolyticum, Ammonifex degensii, Halobacteroides halobius*	integrase in multiple Bacillaceae spp

*refers to hypothetical protein

SUPPLEMENTARY MATERIAL FOR CHAPTER 8 "DISCRIMINATIVE INTERFERENCE MEDIATED BY THE ARGONAUTE PROTEIN IN *T. thermophilus* HORIZONTAL GENE TRANSFER"


Tsco -----MQG----MQVRLNRFILRSLQEEELRPGFWEVRWDPLPSRQENLFQLLGKAANRV
TCCB -----MIGPVQTVEVFLNRFPLRPLEEEEASPWLLQVQLDPPPEPEA-TYALLGRLARQV
Tfilif MNRSKQIGS--KVPVYLNRFPLRPLSPDELRPRLFRPELSPPPDREE-AHPLLAQVARQL
Tth_HB27 ---MNHLG---KTEVFLNRFALRPLNPEELRPWRLEVVLDPPPGREE-VYPLLAQVARRA
Tth_HB8 ---MNHLG---KTEVFLNRFALRPLNPEELRPWRLEVVLDPPPGREE-VYPLLAQVARRA
TSG0.5 ---MNHLG---KTEVFLNRFALRPLNPEELRPWRLEVVLDPPPGREE-VYPLLAQVARRA
Tth_RL ---MNHLG---KTEVFLNRFALRPLNPEELRPWRLEVVLDPPPGREE-VYPLLAQVARRA
T_sp. ---MNYLG---KTEVFLNRFALRPLNPEELRPWRLEVVLDPPPGREE-VYPLLAQVARRA
Tth_JL18 ---MNHLG---KTEVFLNRFALRPLNPEELRPWRLEVVLDPPPGREE-VYPLLAQVARRA
T_NMX2.A1 ---MNHLG---KTEVFLNRFALRPLNPEELRPWRLEVVLDPPPGREE-VYPLLAQVARRA
GGVATPW-QSGLVAWTAPR--IQEGEVRGLSQAYRFTLRDGGALLLDPARVQEREVLSRL
GPMVARF-GTGLVSWSLPEGLALEGQVAGSRGSHAYRLLPQGRRPLDPARPGDREALNAL
GVPAIPWQGAFLLTWEEPK--RMQGEVRRGERVYAFSLEAQGRAELDPARPQDREALSRL
GGVTVRM-GDGLASWSPPEVLVLEGTLARMGQTYAYRLYPKGRRPLDPKDPGERSVLSAL
GGVTVRM-GDGLASWSPPEVLVLEGTLARMGQNYAYRLYPKGRRPLDPKDPGERSVLSAL
GGVTVRM-GDGLASWSPPEVLVLEGTLARMGQTYAYRLYPKGRKPLDPKDPGERSVLSAL
GGVTVRM-GDGLASWSPPEVLVLEGTLARMGQTYAYRLYPKGRXPLDPKDPGERSVLSAL
GGVTVRM-GDGLASWSPPEVLVLEGTLARMGQTYAYRLYPKGRRPLDPKDPGERSALSSL
GGVTVRM-GDGLASWSPPEVLVLEGTLARMGQTYAYRLYPKGRRPLDPKDPGERSALSSL
GGVTVRM-GDGLASWSPPEVLVLEGTLARMGQTYAYRLYPKGRRPLDPKDPGERSALSSL
AQKKLEAGLKQVYGQS--RRYQVEGNLILGGEIRS-GNGWRIRKGSRLRARVDHTGQLVL
ARRRLELGLQQAFP----KGYRVEGRRLLGLKPLREGPGWRLLRGAELDLMVDPEGRWLL
ASRRLEERLARLRLRRDMQDMNVEERRIYRKEVAS-GPGWRFLQGAELDLLVDQKGRLVL
ARRLLQERLRRL------EGVWVEGLAVYRREHAR-GPGWRVLGGAVLDLWVSDSGAFLL
ARRLLQERLRRL------EGVWVEGLAVYRREHAR-GPGWRVLGGAVLDLWVSDSGAFLL
ARRLLQEHLRRL------KGVWVEGLAVYRREHAR-GPGWRVLGGAVLDLWVSDSGAFLL
ARRLLQERLRRL------EGVWVEGLAVYRREHAR-GPGWRVLGGAVLDLWVSDSGAFLL
ARRLLQERLRRL------EGVWVEGLAVYRREHAR-GPGWRVLGGAVLDLWVSDSGAFLL
ARRLLQERLRRL------EGVWVEGLAVYRREHAR-GPGWRVLGGAVLDLWVSDSGAFLL
ARRLLQERLRRL------EGVWVEGLAVYRREHAR-GPGWRVLGGAVLDLWVSDSGAFLL
EVEIVHRILPTLTLEAWL-QGPPEPRRVRNTYENPPGRRQTWRFLRLEPELKPREVDL-G
EVDLVHRLEATLDLEGWLAQGHPLPARARNAYGGG--YRGVWEVVRLGEE-APEEVLLEN
EVDLFHRILPSLTLEAWLREGYPLPRRARNAYGGP---RRVWDVVRLGEE-APEAVLLPG
EVDPAYRILCEMSLEAWLAQGHPLPKRVRNAYD-----RRTWELLRLGEE-DPKELPLPG
EVDPAYRILCEMSLEAWLAQGHPLPKRVRNAYD-----RRTWELLRLGEE-DPKELPLPG
EVDPAYRILCEMSLEAWLAQGHPLPKRVRNAYD-----RRTWELLRLGEE-DPKELPLPG
EVDPAYRILCEMSLEAWLAQGHPLPKRVRNAYD-----RRTWELLRLGEE-DPKELPLPG
EVDPAYRILCEMSLEAWLAQGHPLPKRVRNAYD-----RRTWELLRLGEE-DPKELPLPG
EVDPAYRILCEMSLEAWLAQGHPLPKRVRNAYD-----RRTWELLRLGEE-DPKELPLPG
EVDPAYRILCEMSLEAWLAQGHPLPKRVRNAYD-----RRTWELLRLGEE-DPKELPLPG
GVSLFDYHAGKGRLSSEEQAGRVVRVAD--DSGAEFFHLTGLLEPVLTLEDLVELGLEAE
GLSLLAYHKQKGHWEGGS-PGRVVWVRDPRRAKE-TPHLTGLLVPVLTLEEL-------E
GKNLLDYHERGRLEG-E-AGQVVWVQDPRRPREEVPHLTGLLLPVLSLEEV----HDLG
GLSLLDYHASKGRLQGRE-GGRVAWVADPKDPRKPIPHLTGLLVPVLTLEDL----HEEE
GLSLLDYHAFKGRLQGRE-GGRVAWVADPKDPRKPIPHLTGLLVPVLTLEDL----HEEE
GLSLLDYHASKGRLQGRE-GGRVAWVADPKDPRKPIPHLTGLLVPVLTLEDL----HEEE
GLSLLDYHASKGRLQGRE-GGRVAWVADPKDPRKPIPHLTGLLVPVLTLEDL----HEEE
GLSLLDYHASKGRLQGRE-GGRVAWVADPKDPRKPIPHLTGLLVPVLTLEDL----HEEE
GLSLLDYHASKGRLQGRE-GGRVAWVADPKDPRKPIPHLTGLLVPVLTLEDL----HEEE
AALALQIGPEERFSQAQKTARGIAERI-FGVP---RAEPVVAQAYLLPKPQLKAGRGQRI
EPLALQIPPEPRRKDAEQVARVVAQRL-PGLP---EARPLRVQARRLPQAQLRAHGGKRV
EVPRLQIPPGERLEGARRTARWVAKHLNLGAPPSLEP--LMDQALRLPLPSLVAAGKRRV
GSLALSLPWEERRRRTREIASWIGRRLGLGTP---EA--VRAQAYRLSIPKLM--GRRAV
GSLALSLPWEERRRRTREIASWIGRRLGLGTP---EA--VRAQAYRLSIPKLM--GRRAV
GSLALSLPWEERRRRTREIASWIGRRLGLGTP---EA--VRAQAYRLSIPKLM--GRRAV
GSLALSLPWEERRRRTREIASWIGRRLGLGTP---EA--VRAQAYRLSIPKLM--GRRAV
GSLALSLPWEERRRRTREIASWIGRRLGLGTP---EA--VRAQAYRLSIPKLM--GRRAV
GGLALSLPWEERRRRTREIASWIGRRLGLGTP---EA--VRAQAYRLSIPKLM--GRRAV
GSLALSLPWEERRRRTREIASWIGRRLGLGTP---EA--VRAQAYRLSIPKLM--GRRAV


Tsco GKPADALRKGALKAQDATVGLLVVEGKPEWPQALRKSLQDMARASGVRLRLADPIAVARG
TCCB GKPADALRVGALGAREATVALLRLDGGQGWPPLLLEALKGVARASGVSMRFL--IGQTPL
Tfilif KKPADALEVGPFRSRPAKVALLRLDGGRGWPPFLERRLERLGLETG---------TLEAD
Tth_HB27 SKPADALRVGFYRAQETALALLRLDGAQGWPEFLRRALLRAFGASGASLRLH---TLHAH
Tth_HB8 SKPADALRVGFYRAQETALALLRLDGAQGWPEFLRRALLRAFGASGASLRLH---TLHAH
TSG0.5 SKPADALRVGLYRAQETALALLRLDGAQGWPEFLRRALLQAFGAGGASLRLH---TLHAH
Tth_RL SKPADALRVGLYRAQETALALLRLDGAQGWPEFLRRALLQAFGASGASLRLH---TLHAH
T_sp. SKPADALRVGLYRAQETALALLRLDGAQGWPEFLRRALLQAFGAGGASLRLH---TLHAH
Tth_JL18 SKPADALRVGLYRAQETALALLRLDGAQGWPEFLRRALLQAFGAGGASLRLH---TLHAH
T_NMX2.A1 SKPADALRVGLYRAQETALALLRLDGAQGWPEFLRRALLQAFGAGGASLRLH---TLHAH
DLRGLDLASSLDQLAQQEVAALLVQTPPLTWEERNRLKRASLQRTLPSQFFNPPLAE---
TCCB PERELAFRQALEGIRAQGAEALLVLTPPLSPERRNRLKALALQEGLPTQLLNTPLGP--G
Tfilif LQDELALRRKLSALREAGFSALLVLTPPLSWEVRNRLKALCLKEGLPTQLLNTPLKEAKE
Tth_HB27 PSQGLAFREALRKAKEEGVQAVLVLTPPMAWEDRNRLKALLLREGLPSQILNVPLRE--E
Tth_HB8 PSQGLAFREALRKAKEEGVQAVLVLTPPMAWEDRNRLKALLLREGLPSQILNVPLRE--E
TSG0.5 PSQGLAFREALRKAKEKGVQAVLVLTPPMAWEDRNRLKALLLREGLPSQILNVPLRE--E
Tth_RL PSQGLAFREALRKAKEKGVQAVLVLTPPMAWEDRNRLKALLLREGLPSQILNVPLRE--E
T_sp. PSQGLAFREALRKAKEKGVQAVLVLTPPMAWEDRNRLKALLLREGLPSQILNVPLRE--E
Tth_JL18 PSQGLAFREALRKAKEKGVQAVLVLTPPMAWEDRNRLKALLLREGLPSQILNVPLRE--E
T_NMX2.A1 PSQGLAFREALRKAKEKGVQAVLVLTPPMAWEDRNRLKALLLREGLPSQILNVPLRE--E
--HKLDNILLGLASKLGWRMVGLEGEYPAEVAVGEDAGVDGSRSLRYGGAACAVTADGGL
TCCB DRHRMANALLGLLAKLGWPVVALEGRYAAELAVGEDAG--GRDTLRFGGAACATTGDGSF
Tfilif EEHRLANALLGLLVKAGFQLLALEGEYPAELVVGEDAG--GQNSFRFGGAACAVGQDGSL
Tth_HB27 ERHRWENALLGLLAKAGLQVVALSGAYPAELAVGEDAG--GRESFRFGGAACAVGGDGGH
Tth_HB8 ERHRWENALLGLLAKAGLQVVALSGAYPAELAVGEDAG--GRESFRFGGAACAVGGDGGH
TSG0.5 ERHRWENALLGLLAKAGLQVVALSGAYPAELAVGEDAG--GRESFRFGGAACAVGGDGGH
Tth_RL ERHRWENALLGLLAKAGLQVVALSGAYPAELAVGEDAG--GRESFRFGGAACAVGGDGGH
T_sp. ERHRWENALLGLLAKAGLQVVALSGAYPAELAVGEDAG--GRESFRFGGAACAVGGDGGH
Tth_JL18 ERHRWENALLGLLAKAGLQVVALSGAYPAELAVGEDAG--GRESFRFGGAACAVGGDGGH
T_NMX2.A1 ERHRWENALLGLLAKAGLQVVALSGAYPAELAVGEDAG--GRESFRFGGAACAVGGDGGH
LSWILPEAQRGERIHEDVVWGMVQEVLLRFRRLAGRWPRHVLLLRDGKTQRQEFTLALEE
TCCB LGWALPEAQVGERIPERVAWDLLANTLKGFHNAYGRFFRHVLLLRDGRIPQGEFRLALEG
Tfilif LTWALPEAQPGERIPEEVVWDLLQEALLAFSPE-GRLSRKVLLLRDGKVPPGEFARALEE
Tth_HB27 LLWTLPEAQAGERIPQEVVWDLLEETLWAFRRKAGRLPSRVLLLRDGRVPQDEFALALEA
Tth_HB8 LLWTLPEAQAGERIPQEVVWDLLEETLWAFRRKAGRLPSRVLLLRDGRVPQDEFALALEA
TSG0.5 LLWTLPEAQAGERIPQEVVWDLLEETLWAFRRKAGRLPSRVLLLRDGRVPQDEFALALEA
Tth_RL LLWTLPEAQAGERIPQEVVWDLLEETLWAFRRKAGRLPSRVLLLRDGRVPQDEFALALEA
T_sp. LLWTLPEAQAGERIPQEVVWDLLEETLWAFRRKAGRLPSRVLLLRDGRVPQDEFALALEA
Tth_JL18 LLWTLPEAQAGERIPQEVVWDLLEETLWAFRRKAGRLPSRVLLLRDGRVPQDEFALALEA
T_NMX2.A1 LLWTLPEAQAGERIPQEVVWDLLEETLWAFRRKAGRLPSRVLLLRDGRVPQDEFALALEA
LARKEIGYDLVSVRKSSGGRVYPVEGGRLRDGLYVPLPEEEDKVSFLLLTAYPGEGRMKG
TCCB LRQEEISYDLISVRKTGGGRIYPVR-GRLGDGVFLPLEGD----AFLLLTVHRE---GRG
Tfilif LRRRGVAYDLLSVRKSGGGRLYPVR-GRLADGLFLPLEEE----TFLLLTVHRE---GRG
Tth_HB27 LAREGIAYDLVSVRKSGGGRVYPVQ-GRLADGLYVPLEDK----TFLLLTVHRD---FRG
Tth_HB8 LAREGIAYDLVSVRKSGGGRVYPVQ-GRLADGLYVPLEDK----TFLLLTVHRD---FRG
TSG0.5 LAREGIPYDLVSVRKSGGGRVYPVQ-GRLADGLYVPLEDR----TFLLLTVHRD---FRG
Tth_RL LAREGIAYDLVSVRKSGGGRVYPVQ-GRLADGLYVPLEDR----TFLLLTVHRD---FRG
T_sp. LAREGIAYDLVSVRKSGGGRVYPVQ-GRLADGLYVPLEDR----TFLLLTVHRD---FRG
Tth_JL18 LAREGIAYDLVSVRKSGGGRVYPVQ-GRLADGLYVPLEDR----TFLLLTVHRD---FRG
T_NMX2.A1 LAREGIAYDLVSVRKSGGGRVYPVQ-GRLADGLYVPLEDR----TFLLLTVHRD---FRG
TPRPLKVVHEEGPTPVEELARQIYHLSRLYPPSGYRFPSLPAPLHLADRLVREVGRVGLS
TCCB TPRPLKVVREAGETPLEEVARQVYHLTRLYPPSGFLFPRLPAPLHLADRLVKERARVGSF
Tfilif TPRPLKLVREEGNTPLVDLALQIYHLTRLYPASGFLFPRLPAPLHLADRLVREVGRLGVR
Tth_HB27 TPRPLKLVHEAGDTPLEALAHQIFHLTRLYPASGFAFPRLPAPLHLADRLVKEVGRLGIR
Tth_HB8 TPRPLKLVHEAGDTPLEALAHQIFHLTRLYPASGFAFPRLPAPLHLADRLVKEVGRLGIR
TSG0.5 TPRPLKLVHEAGDTPLEALAHQIFHLTRLYPASGFAFPRLPAPLHLADRLVKEVGRLGIR
Tth_RL TPRPLKLVHEAGDTPLEALAHQIFHLTRLYPASGFAFPRLPAPLHLADRLVKEVGRLGIR
T_sp. TPRPLKLVHEAGDTPLEALAHQIFHLTRLYPASGFAFPRLPAPLHLADRLVKEVGRLGIR
Tth_JL18 TPRPLKLVHEAGDTPLEALAHQIFHLTRLYPASGFAFPRLPAPLHLADRLVKEVGRLGIR
T_NMX2.A1 TPRPLKLVHEAGDTPLEALAHQIFHLTRLYPASGFAFPRLPAPLHLADRLVKEVGRLGIR
SLHGLDREKLFFV
TCCB PQGRIQPSSLFFV
Tfilif HLKEVPRDRLFFV
Tth_HB27 HLKEVDREKLFFV
Tth_HB8 HLKEVDREKLFFV
TSG0.5 HLKEVDREKLFFV
Tth_RL HLKEVDREKLFFV
T_sp. HLKEVDREKLFFV
Tth_JL18 HLKEVDREKLFFV
T_NMX2.A1 HLKEVDREKLFFV

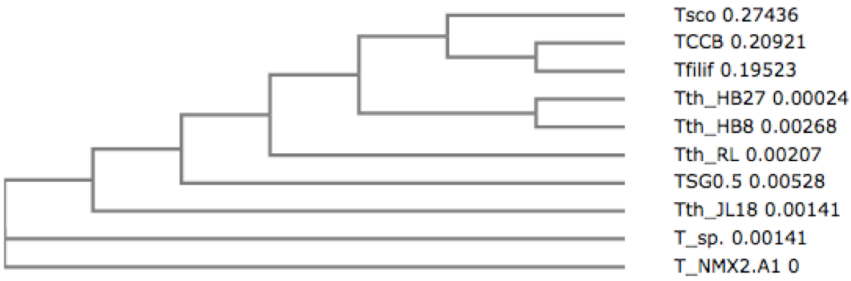


FIGURE S8.1. Sequence alignment of TtAgo in different *Thermus* strains. Argonaute homologues to TtAgo from 10 different species were sequence-aligned: Tsco (*T. scotoductus SA01*), TCCB (*Thermus sp. CCB US3 UF1*), Tfilif (*T. filiformis ATCC 43280*), Tth_HB27 (*T. thermophilus HB27*), Tth_HB8 (*T. thermophilus HB8*), TSG0.5 (*T. thermophilus SG0.5JP17-16*), Tth_RL (*Thermus sp. RL*), T_sp. (*Thermus* sp. Isolate 2.9), Tth_JL18 (*T. thermophilus JL18*), and T_NMX2.A1 (*Thermus sp. NMX2.A1*). The predicted sketch of the phylogenetic tree (based on neighbourhood-joining) of these aligned sequences is also provided, indicating the length of the branch after each name. Catalytic residues involved in DNA cleavage are framed in black (D478, E512, D546 and D660).

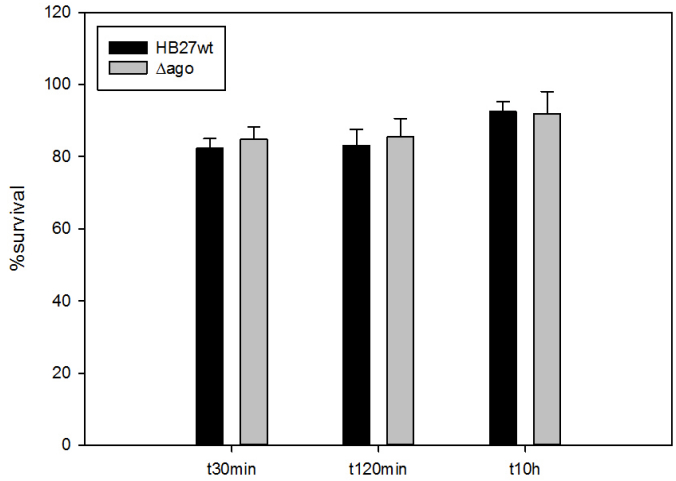


FIGURE S8.2. Environmental stress assays on *Δago strains*. Heat shock tests, measured as percentage of cell survival, estimated following the equation: *survival = log (CFU_{ti})/ log (CFU_{t0}) *100*; where CFU_{ti} refers to the CFU grown in TB media *i* minutes after exposure to 79.2 ºC, and CFU_{t0} refers to the CFU before such temperature raise. Cells from HB27wt (*ago*$^{+}$, black bars) and *Δago* (*ago*$^{-}$, grey bars) were grown at 60 ºC until OD_{550}= 0.8. Then, cells were rapidly transferred to an incubator at 79.2 ºC for further growth and platted after different timings after the heat shock (30 minutes (t30 min), 120 minutes (t120min) and 10 hours (t10h)). Error bars represent the mean standard deviation (n=3).

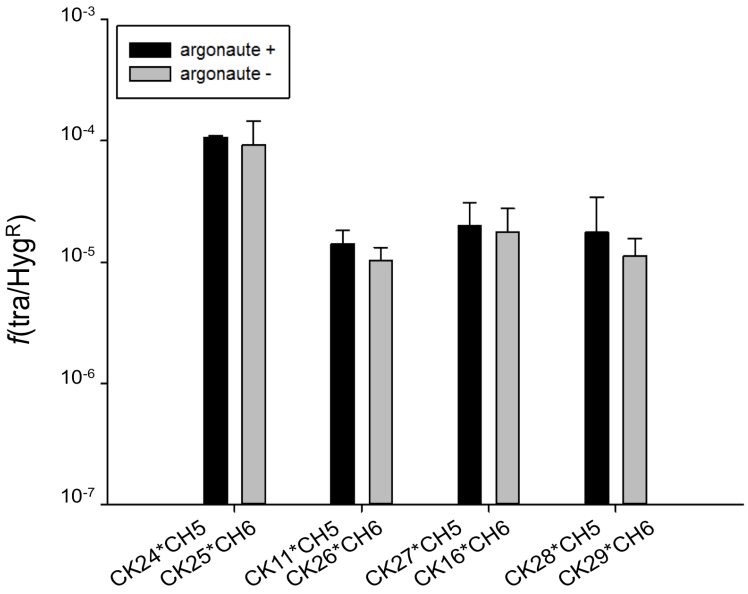


FIGURE S8.3. Conjugation-like DNA transfer does not elicit TtAgo DNA-DNA interference, regardless which loci is tagged. Transfer frequencies obtained after crossing equal cell amounts of CH5 (*Δago*) or CH6 (*ago*$^{+}$) with Kan^{R} mutants labelled in different chromosomal loci (CK11, CK16, CK24-29), that were either *ago*$^{+}$ (grey bars) or *Δago* (black bars). Frequencies represent an average of 5 independent experiments. Error bars correspond to the mean standard deviation. Differences of transfer frequencies between *ago*$^{+}$ and *Δago* mutants were not significant (*p*-value: 0.968). Likewise, no significant difference among loci groups neither within each loci group could be detected (*p*-value: 0.339, *p*-value: 0.105, respectively).

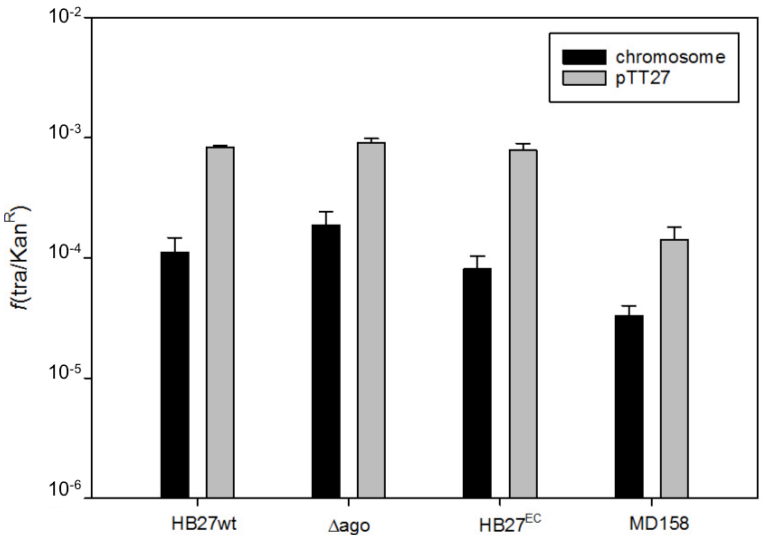


FIGURE S8.4. Matings involving *ago*$^{-}$ derivatives follow the preference for transfer of pTT27-associated genes. Transfer frequencies obtained after crossing equal cell amounts of chromosomal Kan^{R} mutants with Hyg^{R} chromosomally-labelled strains (black bars) and those labelled in the pTT27 megaplasmid (grey bars) from HB27wt (*ago*$^{+}$) derivatives and those *ago*$^{-}$ derivatives from strains *Δago,* HB27^{EC} and MD158. Frequencies represent an average of 3 independent experiments. Error bars correspond to the mean standard deviation. In average, pTT27-associated genes were transferred preferentially to those located in the chromosome in all cases.

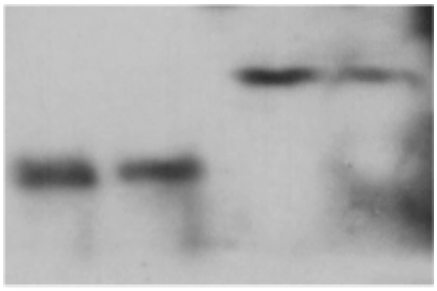


FIGURE S8.5. Western-blot of strains harbouring fluorescent fusion to TtAgo. Immunoblotting detection of sGFP and sYFP reporters fused to the C-terminal region of TtAgo (PH17 and PH18, respectively) using α- TtAgo antiserum diluted 1:3000. First two lanes correspond to HB8 and HB27wt controls, respectively.

ANNEX IX

Results obtained during the development of the PhD are entailed in the following publications:

Alvarez L, Bricio C, Blesa A, Hidalgo A and Berenguer J. (2014). The transferable denitrification of *Thermus thermophilus*. *Applied and Environmental Microbiology*, 80 (1), 19-28

Blesa A, César CE, Averhoff B and Berenguer J. (2015). Noncanonical cell-to-cell DNA transfer in *Thermus* spp is insensitive to Argonaute-mediated interference. *Journal of bacteriology*, 197(1), 138-146

Blesa A and Berenguer J. (2015) Vesicle-protected extracellular DNA contributes to horizontal gene transfer in *Thermus* spp. *International Microbiology*. In press